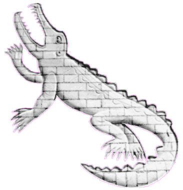
CIC nanoGUNE
nanoscience cooperative research center

Basis sets for SIESTA

Emilio Artacho

Nanogune, Ikerbasque & DIPC, San Sebastian, Spain
Cavendish Laboratory, University of Cambridge



Solving: Basis set

Expand in terms of a finite set of basis functions $\left\{\phi_\mu(\mathbf{r})\right\}:$

$$\psi_n(\mathbf{r}) \approx \sum_\mu \phi_\mu(\mathbf{r})\, c_{\mu,n}$$

unknown

$$\hat{h}\psi_n(\mathbf{r}) = \varepsilon_n \psi_n(\mathbf{r}) \quad \Rightarrow \quad \sum_\mu \left[\hat{h}\,\phi_\mu(\mathbf{r})\right] c_{\mu,n} = \varepsilon_n \sum_\mu \phi_\mu(\mathbf{r})\, c_{\mu,n} \quad \Rightarrow$$

$$\boxed{\sum_\mu h_{\nu\mu} c_{\mu,n} = \varepsilon_n \sum_\mu S_{\nu\mu} c_{\mu,n}}$$

where

$$h_{\nu\mu} \equiv \int \mathrm{d}^3\mathbf{r}\; \phi_\nu^*(\mathbf{r})\, \hat{h}\phi_\mu(\mathbf{r})$$

$$S_{\nu\mu} \equiv \int \mathrm{d}^3\mathbf{r}\; \phi_\nu^*(\mathbf{r})\, \phi_\mu(\mathbf{r})$$

Solving: Basis set

Expand in terms of a finite set of basis functions $\left\{\phi_\mu(\mathbf{r})\right\}:$

$$\psi_n(\mathbf{r}) \approx \sum_\mu \phi_\mu(\mathbf{r})\, c_{\mu,n}$$

$$\rho(\mathbf{r}) = \sum_n^{occ} \left|\psi_n(\mathbf{r})\right|^2 = \sum_n^{occ} \psi_n^*(\mathbf{r})\psi_n(\mathbf{r}) = \sum_{\mu,\nu}\sum_n^{occ} \phi_\mu^*(\mathbf{r})c_{\mu,n}^*\phi_\nu(\mathbf{r})c_{\nu,n} =$$

$$= \sum_{\mu,\nu} \phi_\mu^*(\mathbf{r})\phi_\nu(\mathbf{r}) \boxed{\sum_n^{occ} c_{\mu,n}^* c_{\nu,n}} = \sum_{\mu,\nu} \phi_\mu^*(\mathbf{r})\phi_\nu(\mathbf{r})\, \rho_{\mu\nu}$$

density matrix

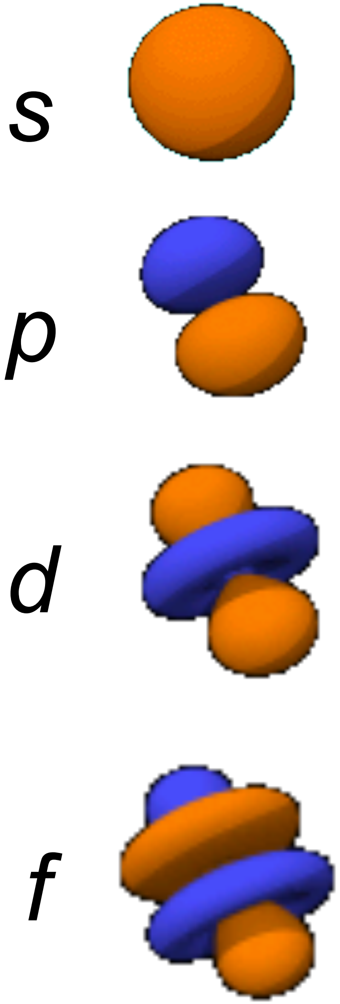
Key matrices

$$h_{\nu\mu} \equiv \int \mathrm{d}^3\mathbf{r}\ \phi_\nu^*(\mathbf{r})\ \hat{h}\phi_\mu(\mathbf{r})$$

$$S_{\nu\mu} \equiv \int \mathrm{d}^3\mathbf{r}\ \phi_\nu^*(\mathbf{r})\ \phi_\mu(\mathbf{r})$$

$$\rho_{\mu\nu} \equiv \sum_n^{occ} c_{\mu,n}^* c_{\nu,n}$$

Basis sets



Plane waves

$$\phi_{\mathbf{G}}(\mathbf{r}) = N\, e^{\mathbf{G}\cdot\mathbf{r}}$$

Atomic orbitals

(There are more…)

Plane wave methods

$$\psi_{i,\vec{k}}(\vec{r}) = \sum_{\vec{g}} c_{i,\vec{g}} \left[\frac{1}{\sqrt{\Omega}} e^{i\left(\vec{k}+\vec{g}\right)\cdot\vec{r}} \right]$$

ADVANTAGES

- Very extended among physicists
- Conceptually simple (Fourier transforms)
- Asymptotically complete
- Allow systematic convergence
- Spatially unbiased (no dependence on the atomic positions)
- “Easy” to implement (FFT)

DISADVANTAGES

- Not suited to represent any function in particular
- Hundreths of wave functions per atom to achieve a good accuracy
- Intrinsic inadequacy for Order-N methods (extended over the whole system)
- Vacuum costs the same as matter
- Hard to converge for tight orbitals (3d …)

M. Payne et al., Rev. Mod. Phys., 64, 1045 (1992)

Atomic orbitals (or atomic-like)

$$\phi_{Ilmn}(\vec{r}) = R_{Iln}(|\vec{r}_I|)\,Y_{lm}(\hat{r}_I)$$

ADVANTAGES

- Very efficient (number of basis functions needed is usually very small).
- Large reduction of CPU time and memory
- Straightforward physical interpretation (population analysis, projected density of states,…)
- Vacuum (almost) for free
- They can achieve very high accuracies…

DISADVANTAGES

- …Lack of systematic for convergence (not unique way of enlarge the basis set)
- Human and computational effort searching for a good basis set before facing a realistic project.
- Depend on the atomic position (Pulay terms appearing in the forces).

Atomic Orbitals: different representations

- Gaussian based + QC machinery

G. Scuseria (GAUSSIAN),
M. Head-Gordon (Q-CHEM)
R. Orlando, R. Dovesi (CRYSTAL)
J. Hutter (CP2K)

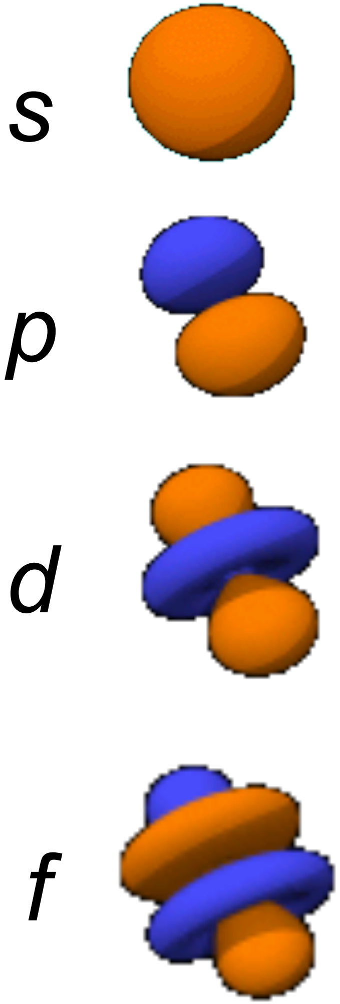
- Slater type orbitals

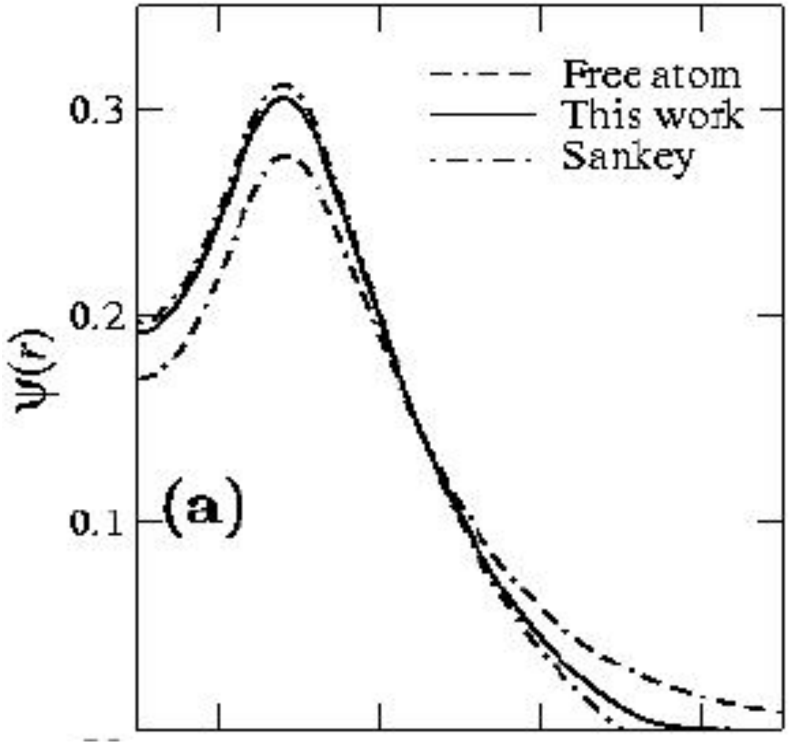
Amsterdam Density Functional

- Numerical atomic orbitals (NAO)

SIESTA
S. Kenny &. A Horsfield (PLATO)
T. Ozaki (OpenMX)
O. Sankey (FIREBALL)

Finite-support atomic orbitals as basis





Strictly localised

(zero beyond cut-off radius)

SIESTA basis sets

The only requirements:

1. $$\phi_{Ilmn}(\vec{r}) = R_{Iln}(|\vec{r}_I|)\,Y_{lm}(\hat{r}_I)$$

2. *Finite support*

They can be:

- *As many as you want (both l-channels and z's)*
- *Of any (radial) shape*
- *Of any cutoff radius*
- *Centred anywhere (not necessarily on atoms)*

SIESTA basis sets

The only requirements:

1. $$\phi_{Ilmn}(\vec{r}) = R_{Iln}(|\vec{r}_I|)\, Y_{lm}(\hat{r}_I)$$

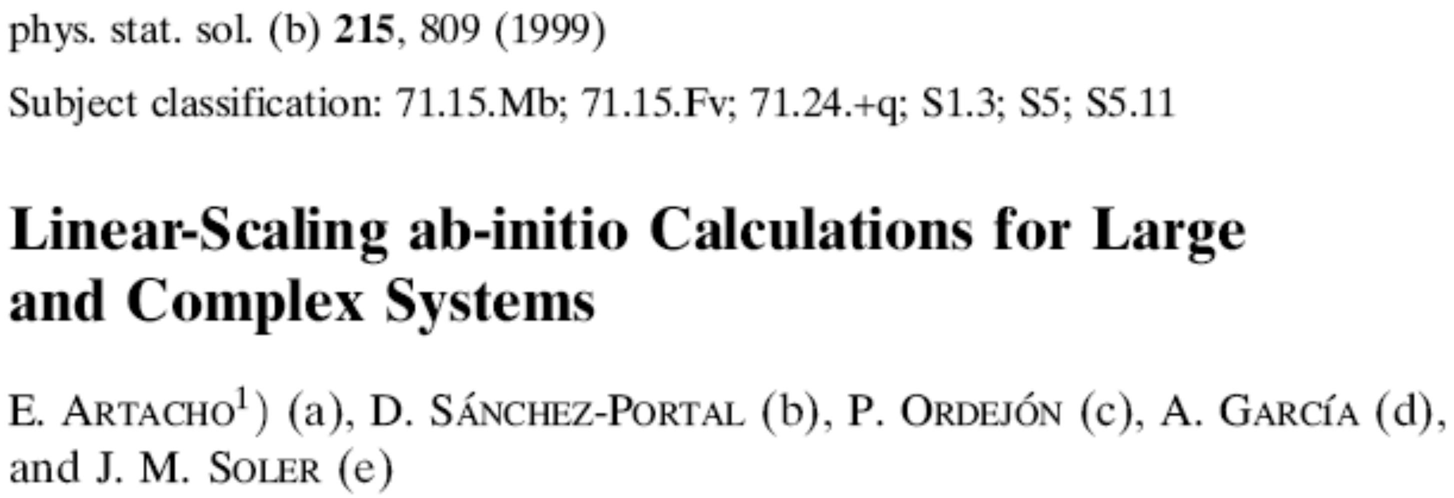
2. *Finite support*

They can be:

- *As many as you want (both l-channels and z's)*
- *Of any (radial) shape*
- *Of any cutoff radius*
- *Centred anywhere (not necessarily on atoms)*

There are NO SIESTA basis sets !!

References

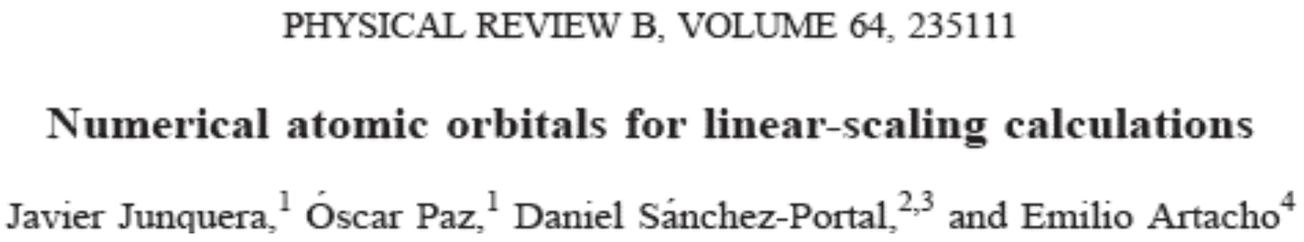


phys. stat. sol. (b) **215**, 809 (1999)

Subject classification: 71.15.Mb; 71.15.Fv; 71.24.+q; S1.3; S5; S5.11

Linear-Scaling ab-initio Calculations for Large and Complex Systems

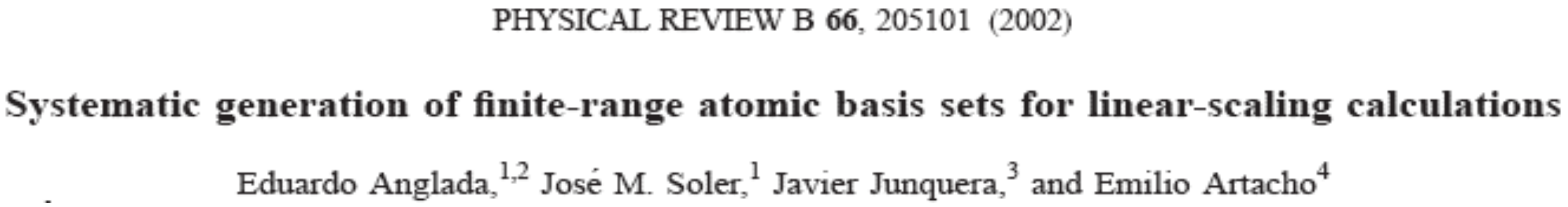
E. ARTACHO[1]) (a), D. SÁNCHEZ-PORTAL (b), P. ORDEJÓN (c), A. GARCÍA (d), and J. M. SOLER (e)



PHYSICAL REVIEW B, VOLUME 64, 235111

Numerical atomic orbitals for linear-scaling calculations

Javier Junquera,[1] Óscar Paz,[1] Daniel Sánchez-Portal,[2,3] and Emilio Artacho[4]



PHYSICAL REVIEW B **66**, 205101 (2002)

Systematic generation of finite-range atomic basis sets for linear-scaling calculations

Eduardo Anglada,[1,2] José M. Soler,[1] Javier Junquera,[3] and Emilio Artacho[4]

Pedestrian guide to basis sets in Quantum Chemistry

- *Minimal basis, or single-z: occupied states (fully or partly) in the free atom*
- *Radial flexibility: multiple zeta (and diffuse orbitals)*
- *Angular flexibility: "polarisation" orbitals*

e.g.

C: Minimal (for the valence): 2s, 2p (2p_x, 2p_y, 2p_z)

Doble-z: two orbitals with different radial shapes for each of the above

Polarisation: add a 3d shell to polarise the 2p shell.

Fe: Minimal: 3d, 4s. Polarisation: 4p to polarise 4s. (4f for 3d)

How to get basis sets for Siesta

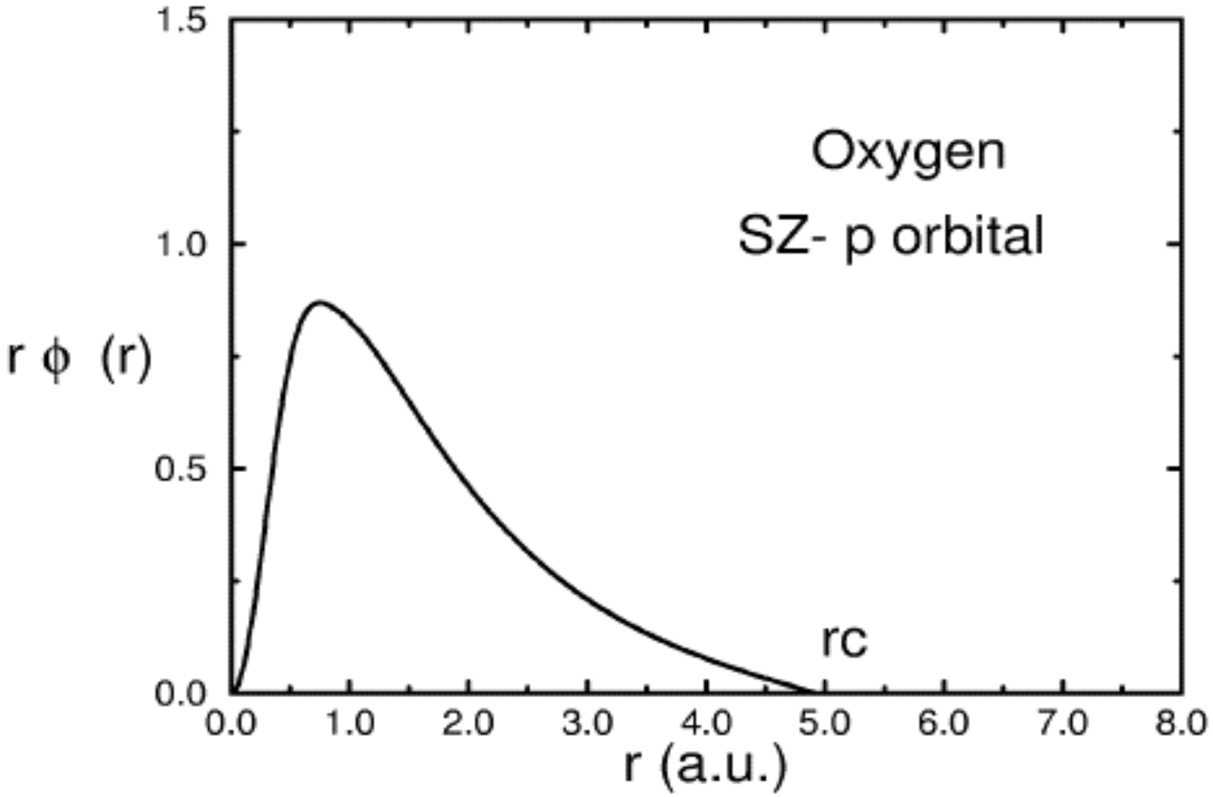
Choice of how many, cutoff radii, and where, made by user.

Radial shapes can also be introduced by user
(Basis type: “user”; a file with a table of values for r (discretised)

Siesta also offers the possibility of generating basis sets:

- *Based on numerical solution of KS DFT on the pseudo-atom + modifications*
- *Quite tunable*
- *Depends on parameters that need to be defined by user*

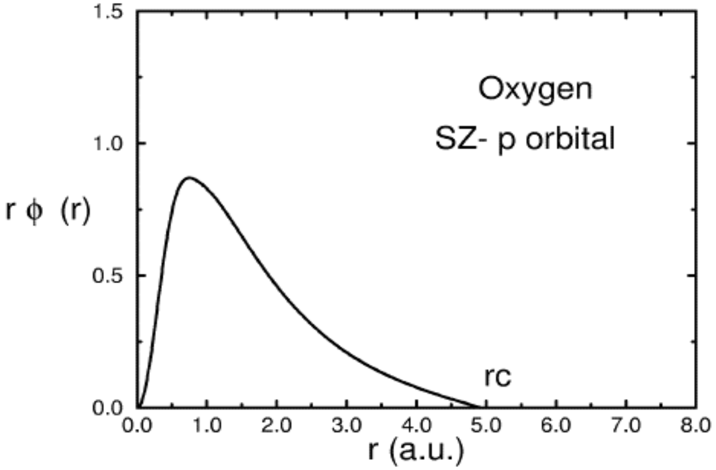
Starting: Minimal basis



Solution of KS-DFT on pseudo-atom,
under an added confinement potential

FIREBALLS O. F. Sankey & D. J. Niklewski, *Phys. Rev. B 40, 3979 (1989)*

Hard confining potentials



Fireballs

O. F. Sankey & D. J. Niklewski,
Phys. Rev. B 40, 3979 (1989)

BUT:

A different cut-off radius for each orbital

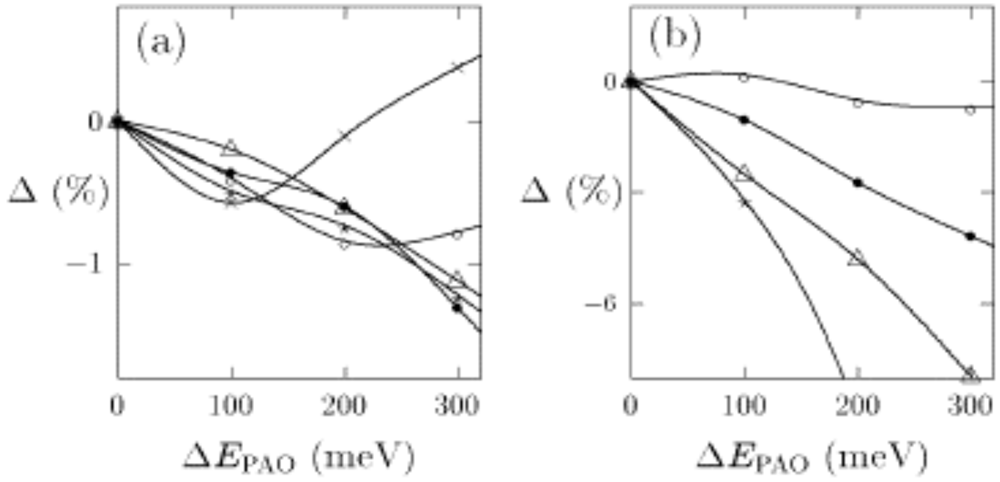
A single parameter

Energy shift

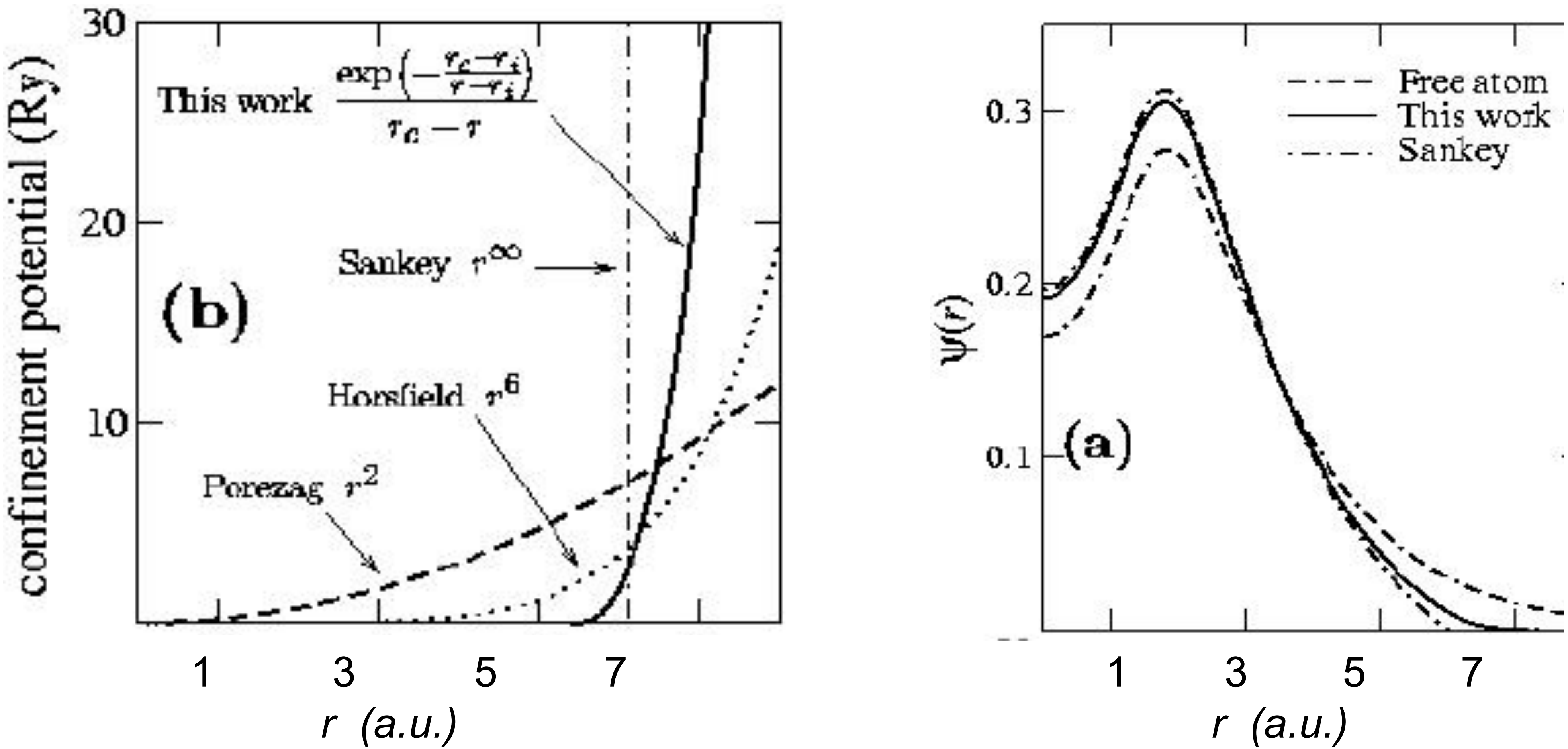
E. Artacho et al. *Phys. Stat. Solidi (b) 215, 809 (1999)*

Convergence vs Energy shift of

Bond lengths *Bond energies*



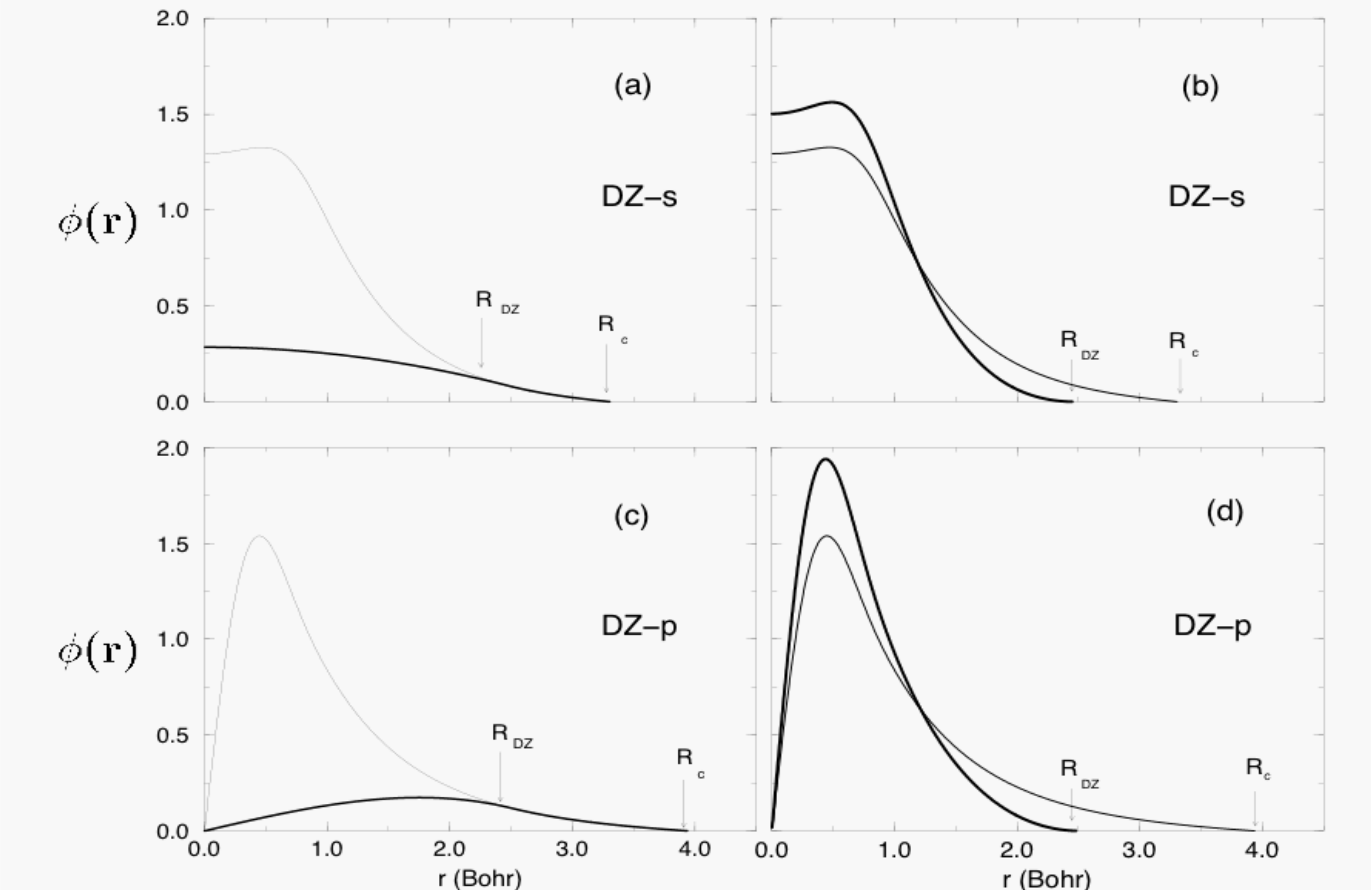
Soft confining potentials



- *Better basis, variationally, & other results*
- *Removes the discontinuity in the derivative*

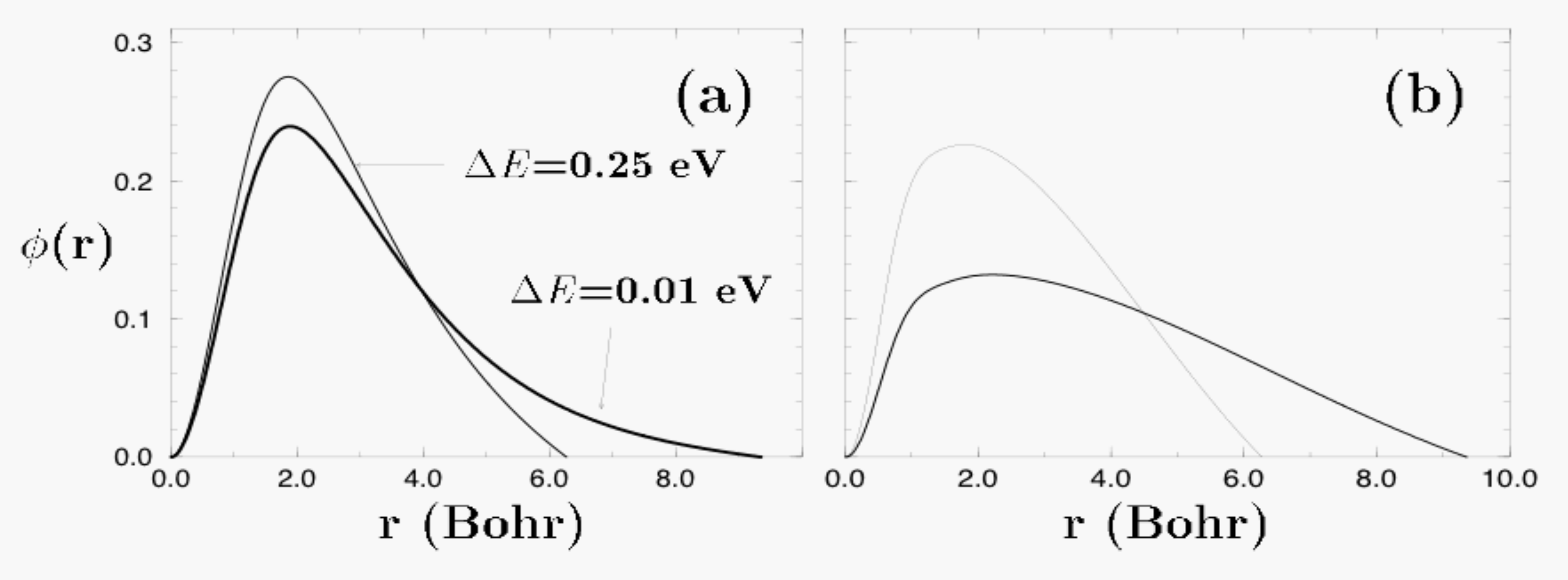
J. Junquera, O. Paz, D. Sanchez-Portal & E. Artacho, *Phys. Rev. B*, **64**, 235111 (2001)
E. Anglada, J. M. Soler, J. Junquera & E. Artacho, *Phys. Rev. B* **66**, 205101 (2002)

Multiple-zeta



E. Artacho *et al.* , *Phys. Stat. Solidi (b)* **215,** 809 (1999).

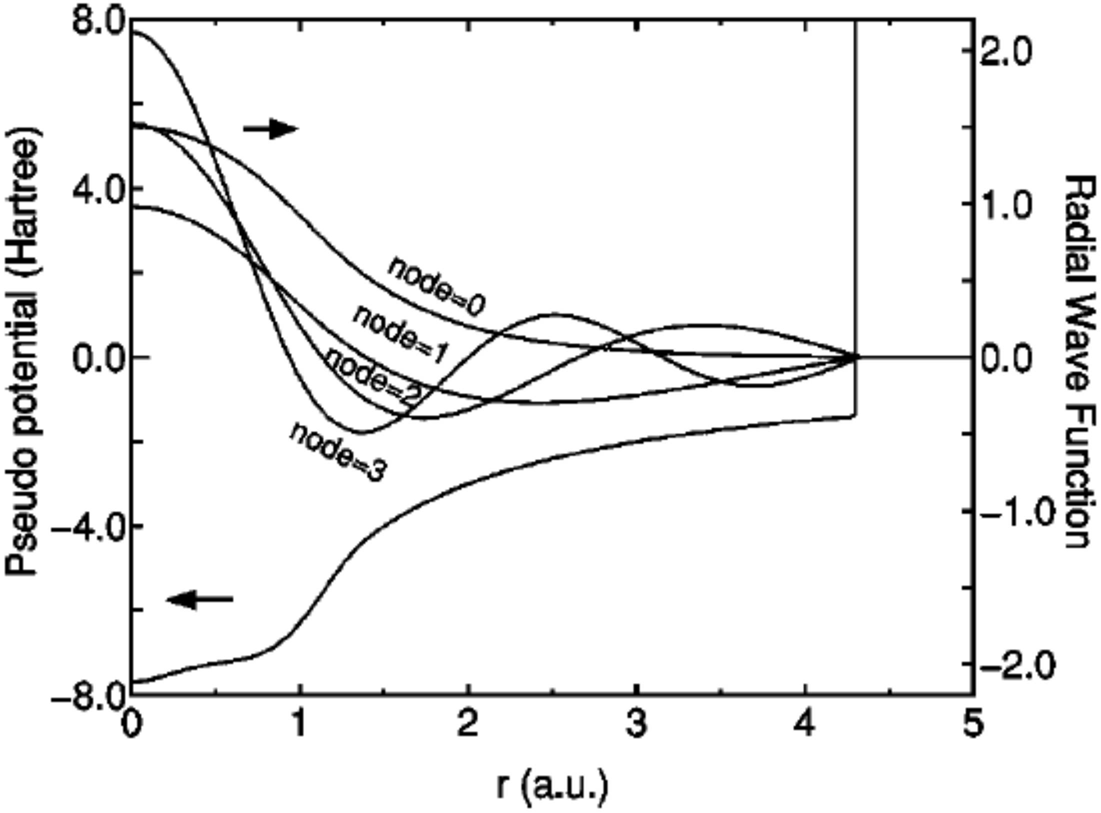
Polarization



E. Artacho *et al.* , *Phys. Stat. Solidi (b)* **215,** 809 (1999).

Schemes to generate multiple-ζ basis sets

Use pseudopotential eigenfunctions with increasing number of nodes



T. Ozaki *et al.*, Phys. Rev. B 69, 195113 (2004)

http://www.openmx-square.org/

Advantages

Orthogonal

Asymptotically complete

Disadvantages

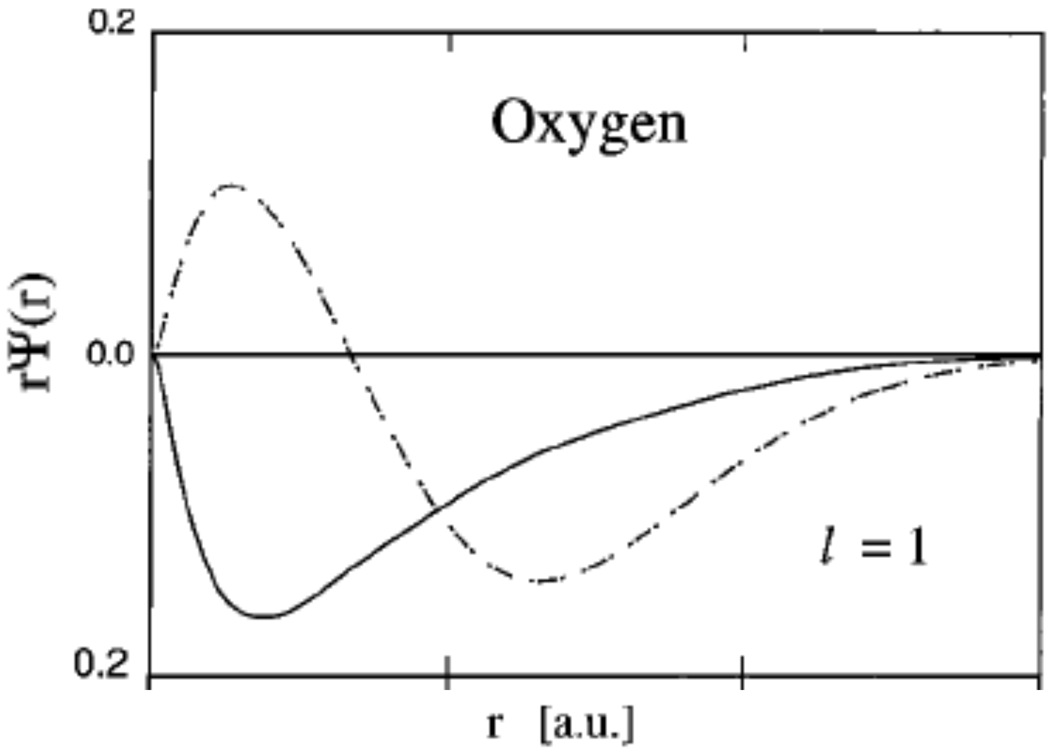
Excited states of the pseudopotentials, usually unbound

Efficient depends on localization ra

Availables in Siesta:

PAO.BasisType   Nodes

Schemes to generate multiple-ζ basis sets
Chemical hardness:
use derivatives with respect to the charge of the atoms



Advantages

Orthogonal

It does not depend on any variational parameter

Disadvantages

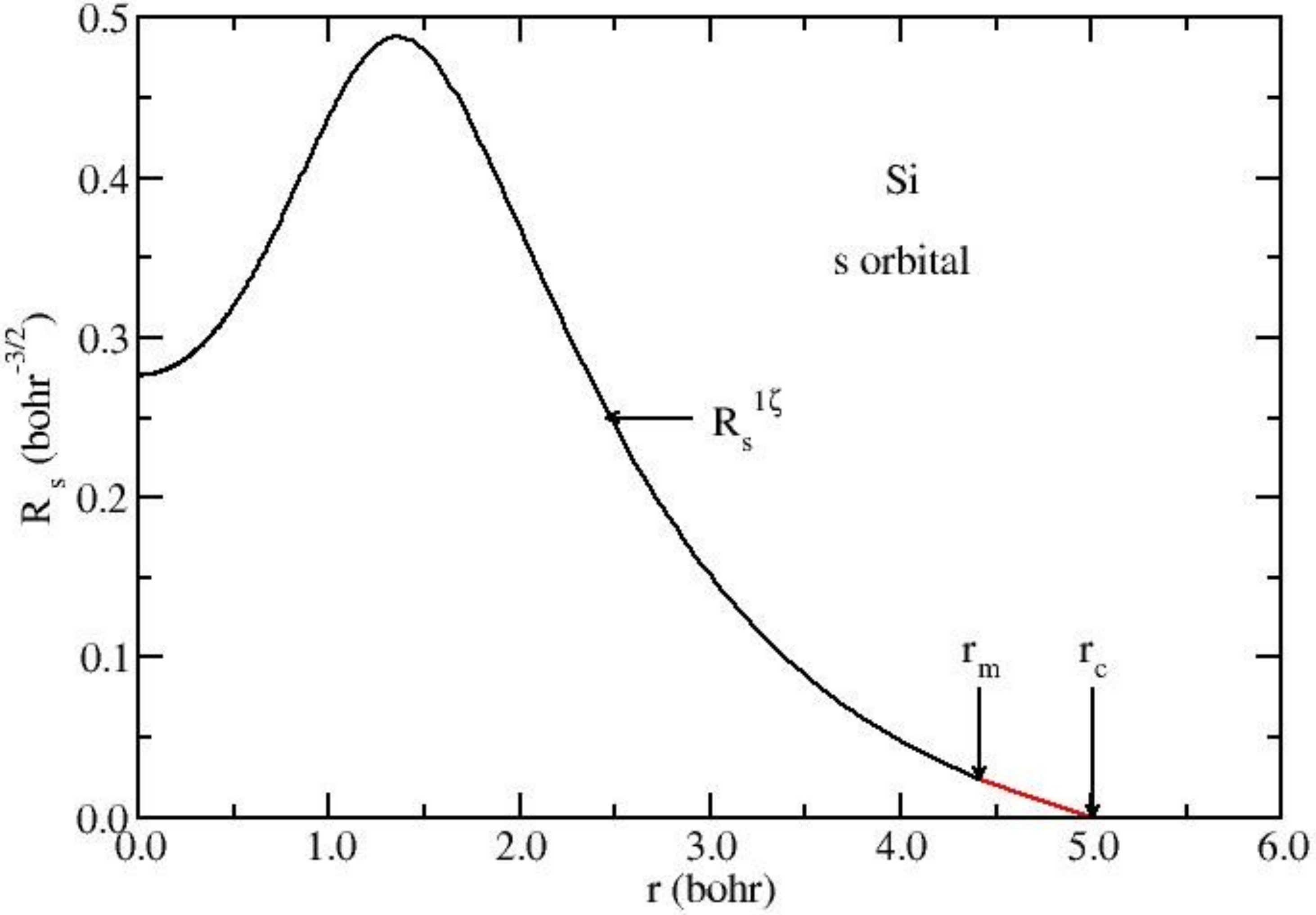
Range of second-ζ equals the rang of the first-ζ function

G. Lippert *et al.*, J. Phys. Chem. 100, 6231 (1996)

http://cp2k.berlios.de/

Default mechanism to generate multiple- ζ in SIESTA:

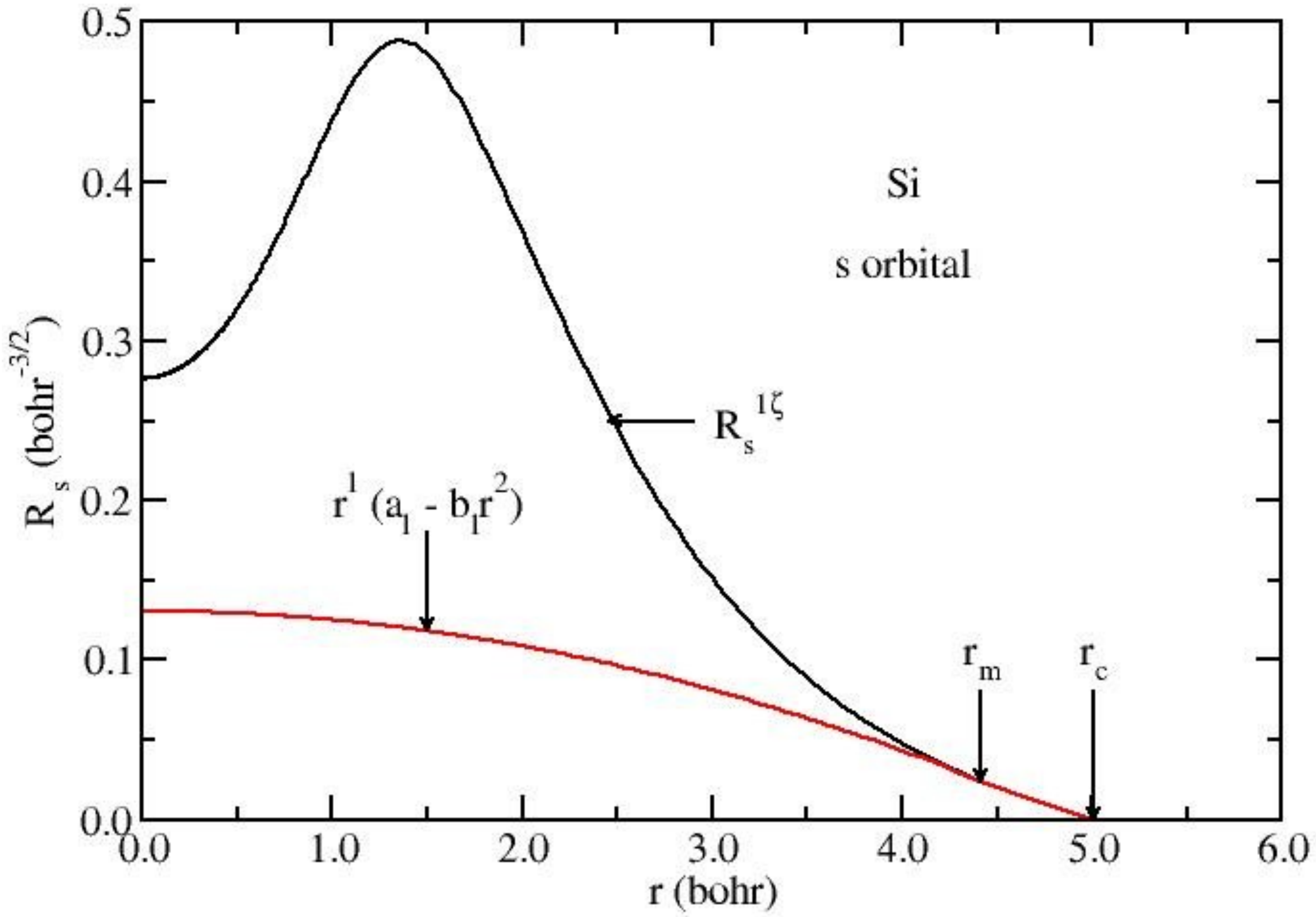
"Split-valence" method



The second-ζ function reproduces the tail of the of the first-ζ outside a radius r_m

Default mechanism to generate multiple- ζ in SIESTA:

"Split-valence" method

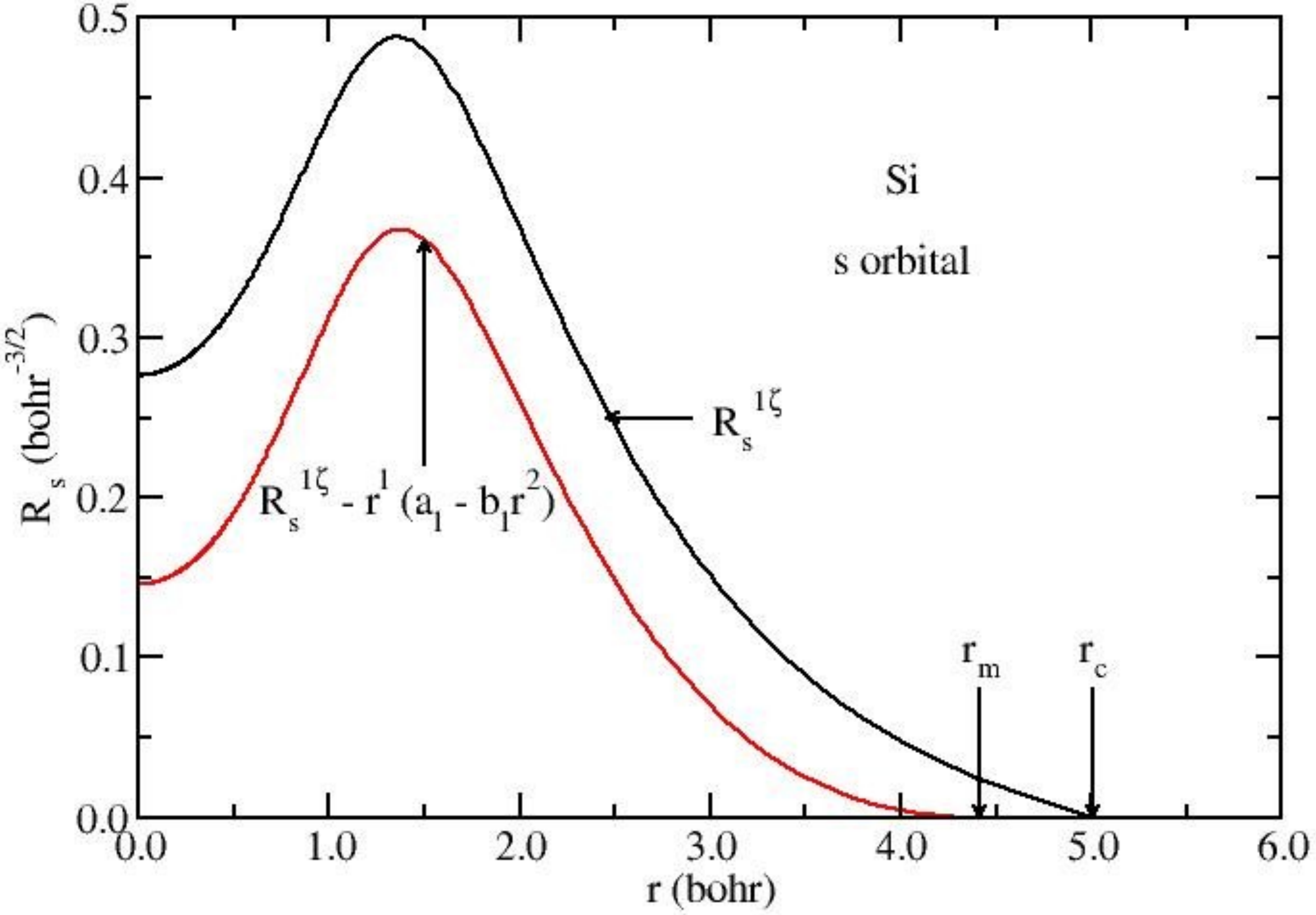


And continuous smoothly towards the origin as $r^l \left(a_l - b_l r^2 \right)$

(two parameters: the second-ζ and its first derivative continuous at r_m

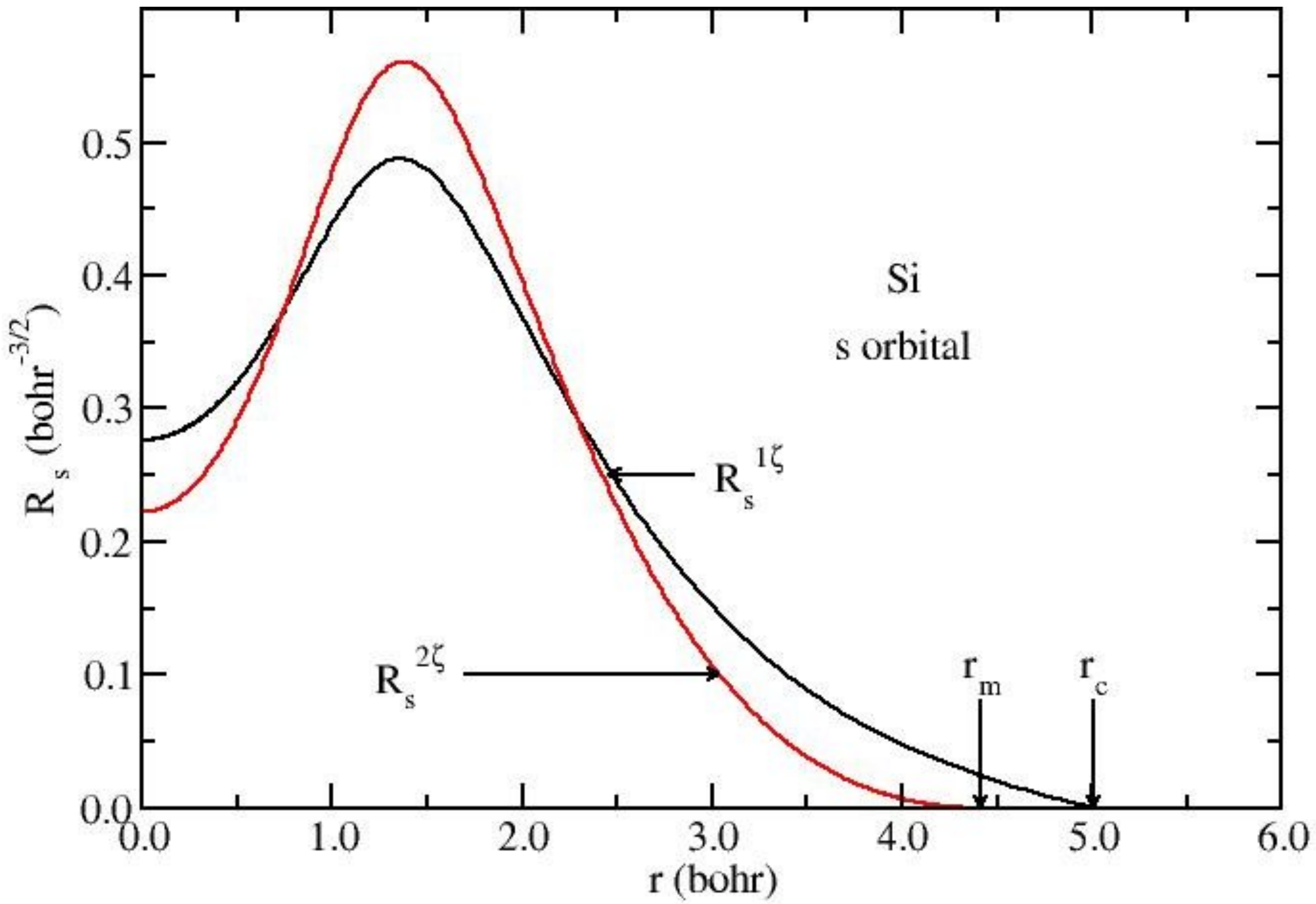
Default mechanism to generate multiple- ζ in SIESTA:

"Split-valence" method



The same Hilbert space can be expanded if we use the difference, with the advantage that now the second-ζ vanishes at r_m (more efficient)

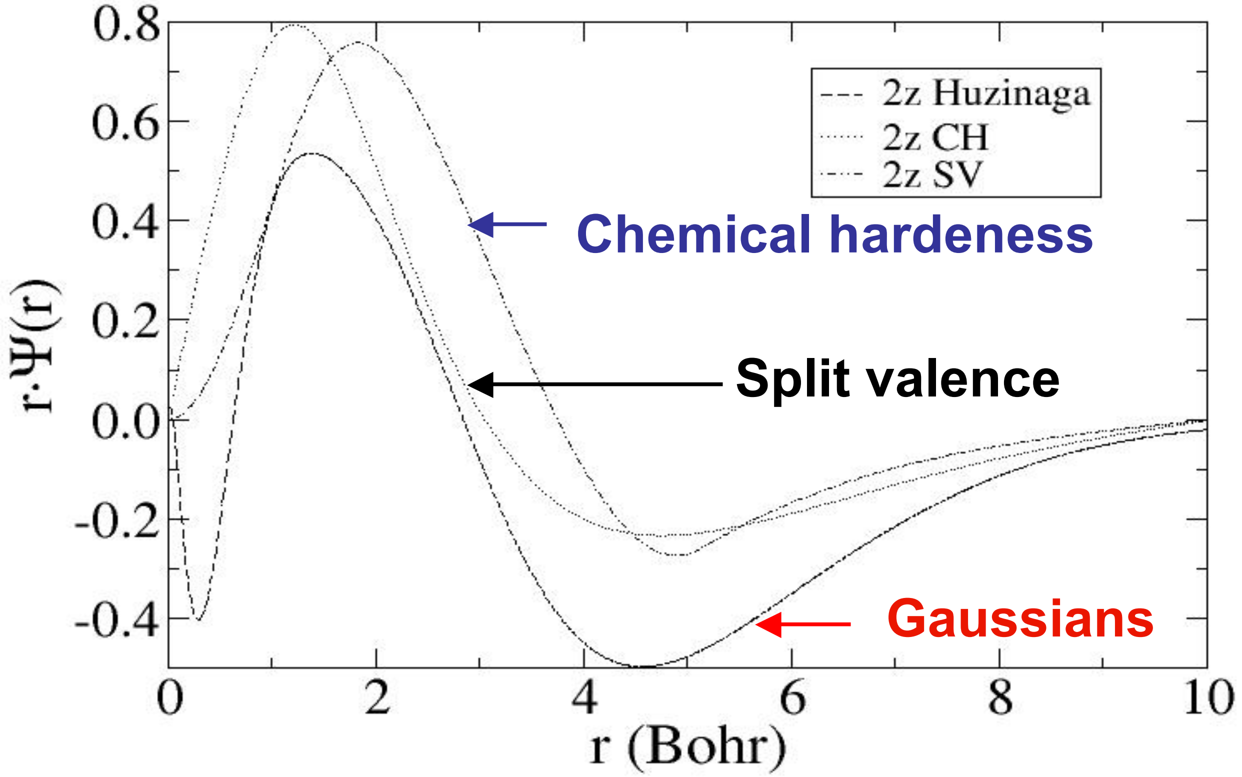
Default mechanism to generate multiple- ζ in SIESTA:
"Split-valence" method



Finally, the second-ζ is normalized

r_m controlled with PAO.SplitNorm (typical value 0.15)

Both split valence and chemical hardness methods provide similar shapes for the second-ζ function



Chemical hardeness

Split valence

Gaussians

Split valence double-ζ has been orthonormalized to firs orbital

SV: higher efficiency

(radius of second-ζ can be restricted to the inner matching radius)

E. Anglada, J. Junquera, J. M. Soler, E. Artacho,

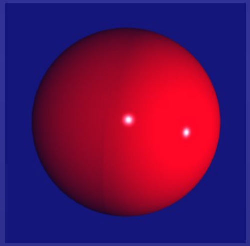
Phys. Rev. B 66, 205101 (2002)

Example of adding angular flexibility to an atom
Polarizing the Si basis set

Si atomic configuration: $\underbrace{1s^2\,2s^2\,2p^6}_{\text{core}}$ $\underbrace{3s^2\,3p^2}_{\text{valence}}$

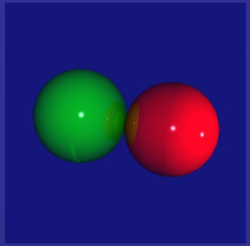
$l = 0$ (s)

$m = 0$

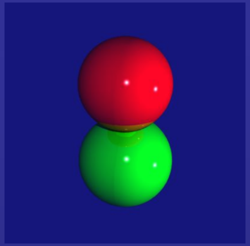


$l = 1$ (p)

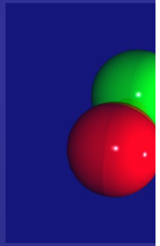
$m = -1$



$m = 0$

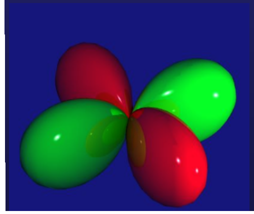


$m =$

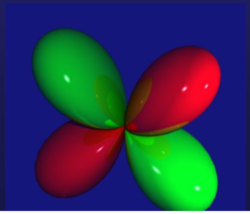


Polarize: add $l = 2$ (d) shell

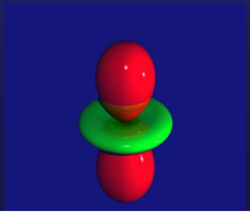
$m = -2$



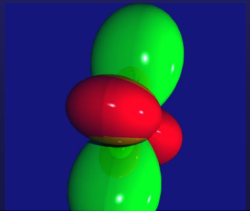
$m = -1$



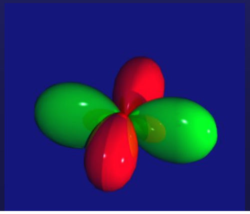
$m = 0$



$m = +1$



$m = +2$

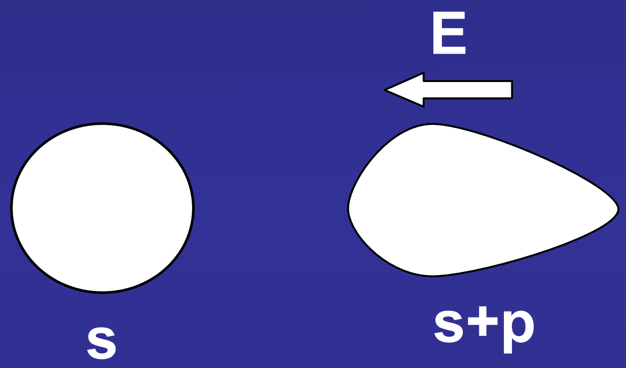


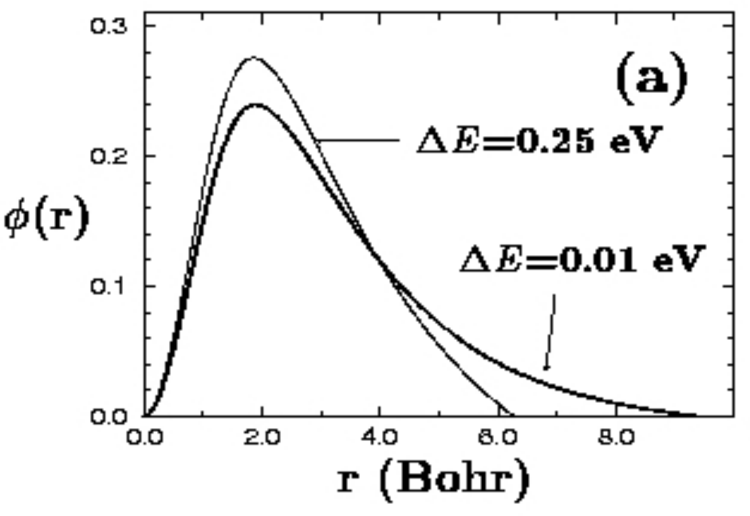
New orbitals directed in different directions with respect the original basi

Two different ways of generate polarization orbitals

Perturbative polarization

Apply a **small electric field** to the orbital we want to polarize





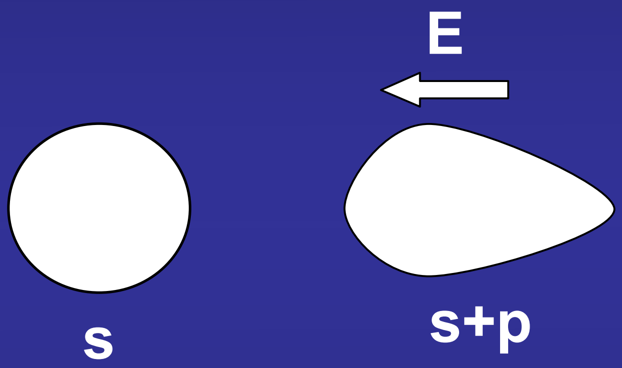
Si 3d orbitals

E. Artacho *et al.* Phys. Stat. Sol. (b), 215, 809 (1999)

Two different ways of generate polarization orbitals

Perturbative polarization

Apply a **small electric field** to the orbital we want to polarize

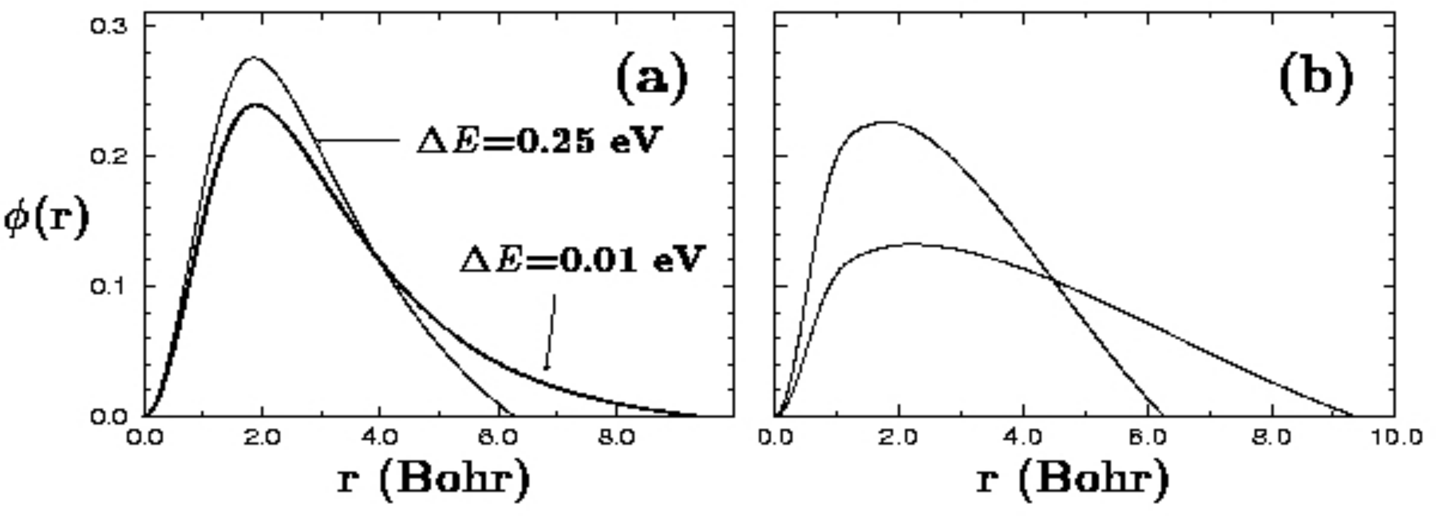


Atomic polarization

Solve **Schrödinger equation** for **higher angular momentum**

unbound in the free atom ⇒

require short cut offs

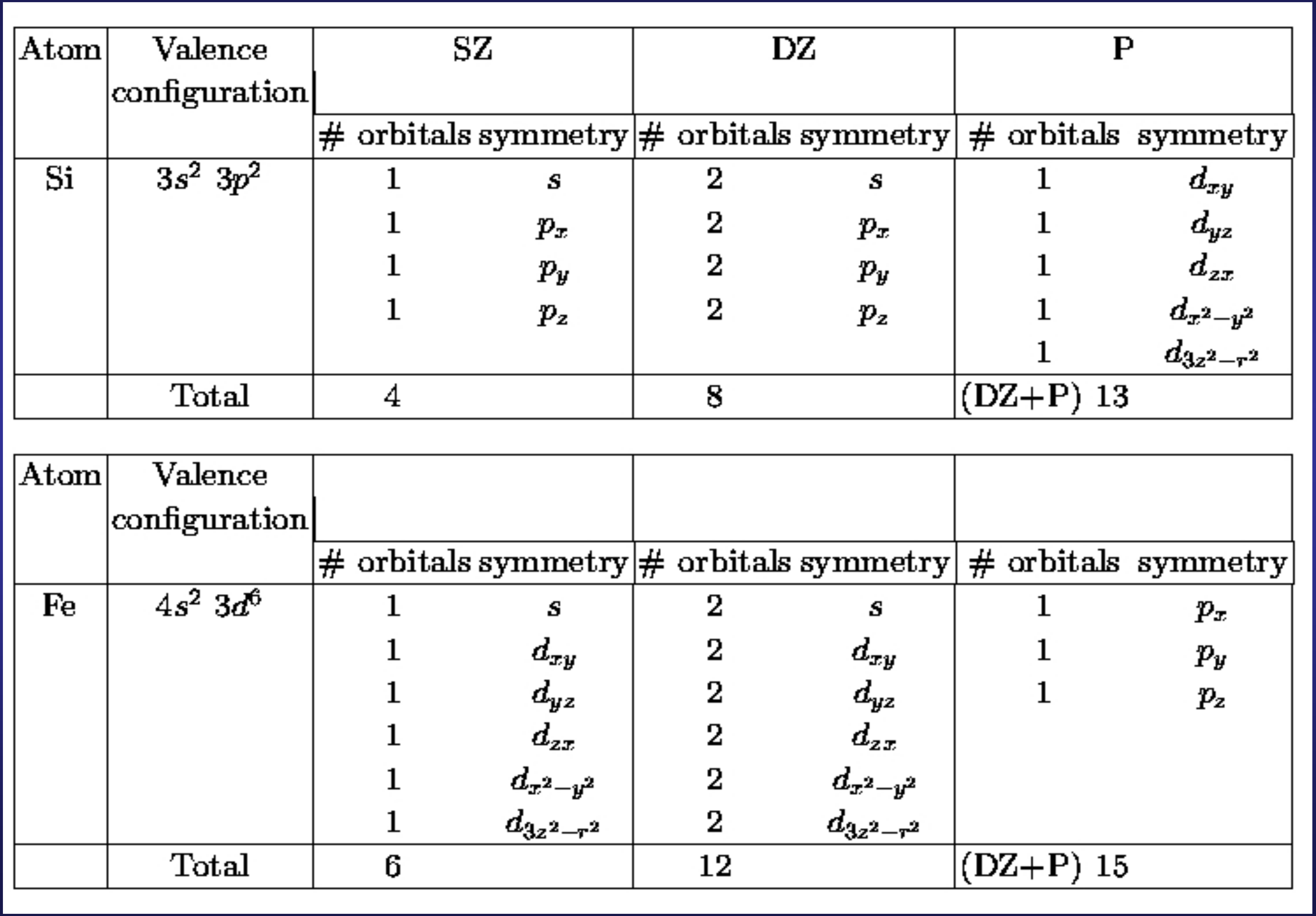


Si 3d

orbitals

E. Artacho *et al.* Phys. Stat. Sol. (b) 215, 809 (1999)

Improving the quality of the basis ⇒ more atomic orbitals per atom



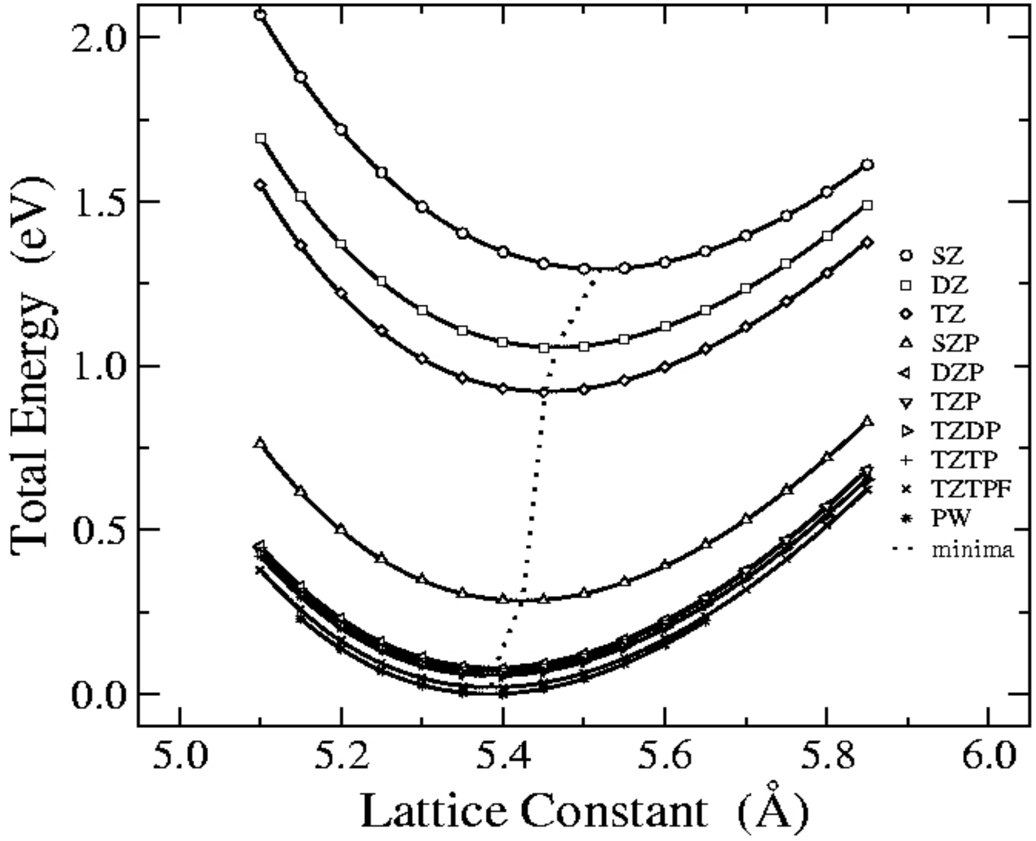
Atom	Valence configuration	SZ		DZ		P	
		# orbitals	symmetry	# orbitals	symmetry	# orbitals	symmetry
Si	$3s^2$ $3p^2$	1	s	2	s	1	d_{xy}
		1	p_x	2	p_x	1	d_{yz}
		1	p_y	2	p_y	1	d_{zx}
		1	p_z	2	p_z	1	$d_{x^2-y^2}$
						1	$d_{3z^2-r^2}$
	Total	4		8		(DZ+P) 13	

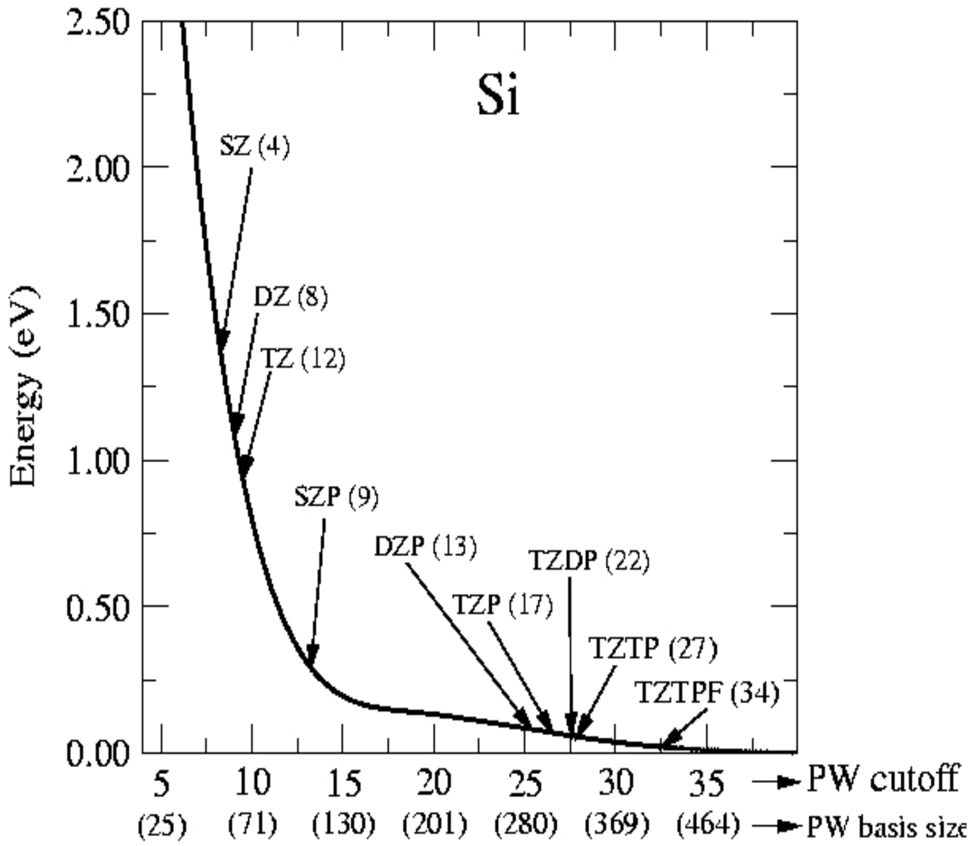
Atom	Valence configuration						
		# orbitals	symmetry	# orbitals	symmetry	# orbitals	symmetry
Fe	$4s^2$ $3d^6$	1	s	2	s	1	p_x
		1	d_{xy}	2	d_{xy}	1	p_y
		1	d_{yz}	2	d_{yz}	1	p_z
		1	d_{zx}	2	d_{zx}		
		1	$d_{x^2-y^2}$	2	$d_{x^2-y^2}$		
		1	$d_{3z^2-r^2}$	2	$d_{3z^2-r^2}$		
	Total	6		12		(DZ+P) 15	

Convergence as a function of the size of the basis se
Bulk Si

Cohesion curves

PW and NAO convergence





Atomic orbitals show nice convergence with respect the size

Polarization orbitals very important for convergence (more than multiple-ζ)

Double-ζ plus polarization equivalent to a PW basis set of 26 Ry

Convergence as a function of the size of the basis se
Bulk Si

	SZ	DZ	TZ	SZP	DZP	TZP	TZDP	PW	APW	Ex
a (Å)	5.52	5.46	5.45	5.42	5.39	5.39	5.39	5.38	5.41	5.4
B (GPa)	89	96	98	98	97	97	96	96	96	98
E_c (eV)	4.72	4.84	4.91	5.23	5.33	5.34	5.34	5.37	5.28	4.6

A DZP basis set introduces the same deviations as the ones due to the DFT or the pseudopotential approaches

SZ = single-ζ

DZ= doble- ζ

TZ=triple- ζ

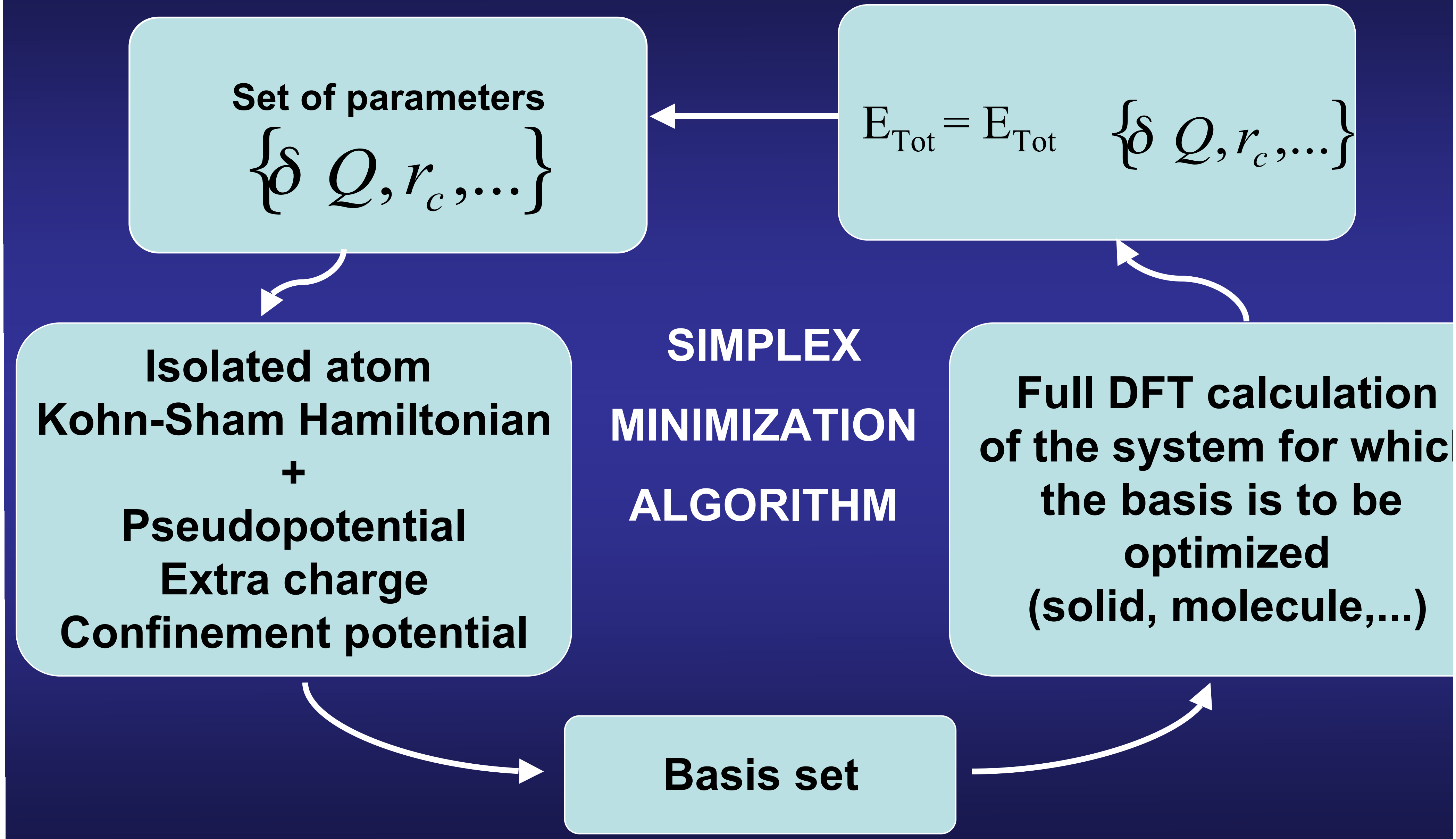
P=Polarized

DP=Doble-polarized

PW: Converged Plane Waves (50 Ry)

APW: Augmented Plane Waves

Optimization of the parameters that define the basis set: the Simplex code



Set of parameters

$$\{\delta Q, r_c, \ldots\}$$

$$\mathrm{E}_{\mathrm{Tot}} = \mathrm{E}_{\mathrm{Tot}} \quad \{\delta Q, r_c, \ldots\}$$

Isolated atom
Kohn-Sham Hamiltonian
+
Pseudopotential
Extra charge
Confinement potential

SIMPLEX
MINIMIZATION
ALGORITHM

Full DFT calculation
of the system for whic
the basis is to be
optimized
(solid, molecule,...)

Basis set

How to introduce the basis set in SIESTA
Effort on defining a systematics with minimum parameters

If nothing is specified: default

Basis size:	PAO.BasisSize	DZP
Range of first-zeta:	PAO.EnergyShift	0.02 Ry
Second-zeta:	PAO.BasisType	Split
Range of second-zeta:	PAO.SplitNorm	0.15
Confinement:	Hard well	

Good basis set in terms of accuracy versus efficiency

More global control on the basis with a few input variables size and range

Size:

Basis size:

PAO.BasisSize	SZ DZ SZP DZP

Range:

Range of first-zeta:	PAO.EnergyShift	0.02 Ry
Range of second-zeta:	PAO.SplitNorm	0.15

The larger both values, the more confined the basis functions

More specific control on the basis: the PAO.Basis block

```
%block PAO.Basis             # Define Basis set
H      1   +0.25             # Species label, number of l-shells, cha
 n=1   0   2                 # n, l, Nzeta
   5.000       3.000         # rc (first-zeta), rm (second-zeta)
   1.000       1.000         # scaling factors
%endblock PAO.Basis
```

More specific control on the basis: the PAO.Basis block

Some variable might be computed automatically

```
%block PAO.Basis                # Define Basis set
H       1    +0.25              # Species label, number of l-shells, cha
 n=1    0    2                  # n, l, Nzeta
   0.000       0.000            # rc (first-zeta), rm (second-zeta)
   1.000       1.000            # scaling factors
%endblock PAO.Basis
```

These variables calculated from

PAO.EnergyShift and PAO.SplitNorm values

More specific control on the basis: the PAO.Basis block

Adding polarization orbitals: perturbative polarization

```
%block PAO.Basis                # Define Basis set
H       1    +0.25              # Species label, number of l-shells, cha
 n=1    0    2 P    1           # n, l, Nzeta, Polarization, NzetaPol
   5.000       3.000            # rc (first-zeta), rm (second-zeta)
   1.000       1.000            # scaling factors
%endblock PAO.Basis
```

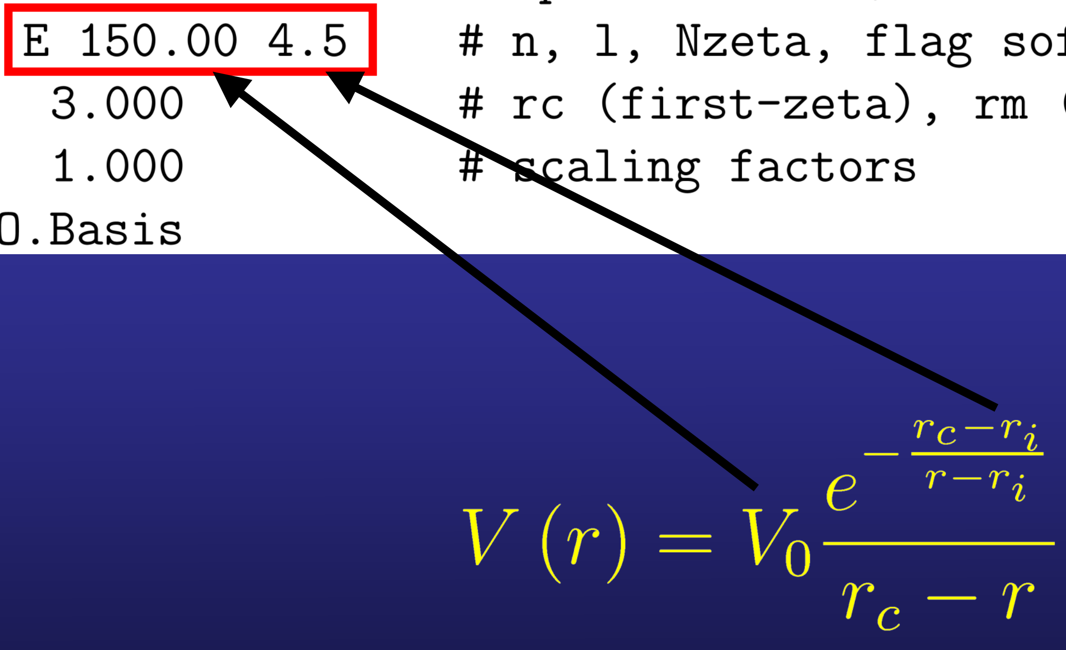
More specific control on the basis: the PAO.Basis block

Adding polarization orbitals: atomic polarization

```
%block PAO.Basis                # Define Basis set
H      2   +0.25                # Species label, number of l-shells, cha
 n=1   0   2                    # n, l, Nzeta
   5.000      3.000             # rc (first-zeta), rm (second-zeta)
   1.000      1.000             # scaling factors
 n=2   1   1                    # n, l, Nzeta
   5.000                        # rc (first-zeta)
   1.000                        # scaling factors
%endblock PAO.Basis
```

More specific control on the basis: the PAO.Basis block

Soft-confinement potential



```
%block PAO.Basis                 # Define Basis set
H       1    +0.25               # Species label, number of l-shells, charge
 n=1    0    2 E 150.00 4.5      # n, l, Nzeta, flag soft-conf, prefactor, inner
   5.000        3.000            # rc (first-zeta), rm (second-zeta)
   1.000        1.000            # scaling factors
%endblock PAO.Basis
```

$$V\left(r\right) = V_0 \frac{e^{-\frac{r_c - r_i}{r - r_i}}}{r_c - r}$$

V_0 in Ry

r_i in bohrs

Recap

Numerical Atomic Orbitals

A very efficient basis set

Spetially suitable for Order-N methods

Smooth transition from quick exploratory calculations to highly converged

Lack of systematic convergence

Current effort for searching the lost systematics. Efficients methods for:

Generate multiple-ζ: Split Valence

Generate polarization orbitals: Perturbative polarization

Control the range of the orbitals in abalanced way: Energy Shift

Confine the orbitals: Soft-confinement potential

A DZP basis set, the same deviations as DFT functional or Pseudo