Max Webinar

BigDFT

Motivations for BigDFT formalism; overview of Daubechies wavelets in DFT

Thierry Deutsch

Laboratoire de Simulation Atomistique - L_Sim

November 12, 2020

Virtual Room

Outline

BigDFT code

- Properties of wavelets
- Precision of the description
- $\mathcal{O}(N)$ BigDFT

Basis sets for DFT calculations

Plane Waves

ABINIT, CPMD, VASP, ...
Systematic convergence

- ✔ Accuracy increases with the number of basis elements
- ✔ Non-localised, optimal for periodic systems
- ✘ Non adaptive

Gaussians, Slater Orbitals

CP2K,Gaussian,AIMPRO, ...
Real space localized

- ✔ Small number of basis functions for moderate accuracy
- ✔ Well suited for molecules and other open structures
- ✘ Non systematic

FFT

Robust, easy to parallelise

Analytic functions

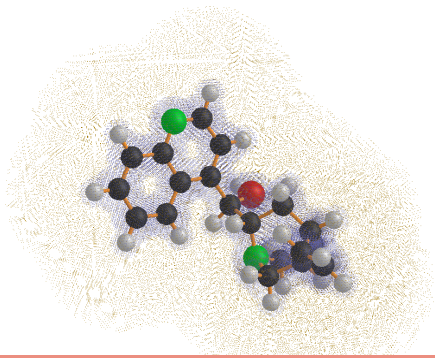
Kinetic and overlap matrices are calculated analytically

Origin of the BigDFT project

STREP European project: BigDFT(2005-2008)

In the beginning: Four partners, 15 people
Now: around 10 active developers, Grenoble, Basel, Barcelona, London, Uppsala, Kobe
Used in production since twelve years.

Aim: To develop an ab-initio DFT code based on Daubechies Wavelets, to be *integrated in ABINIT*. BigDFT 1.0 $\longrightarrow$ January 2008



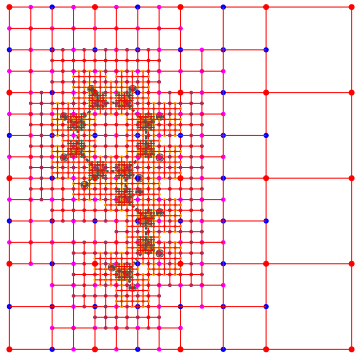
Why have we done this?

- Test the potential advantages of a new formalism
- ☛ A lot of outcomes and interesting results
- Future opportunities and ideas

A DFT code based on Daubechies wavelets

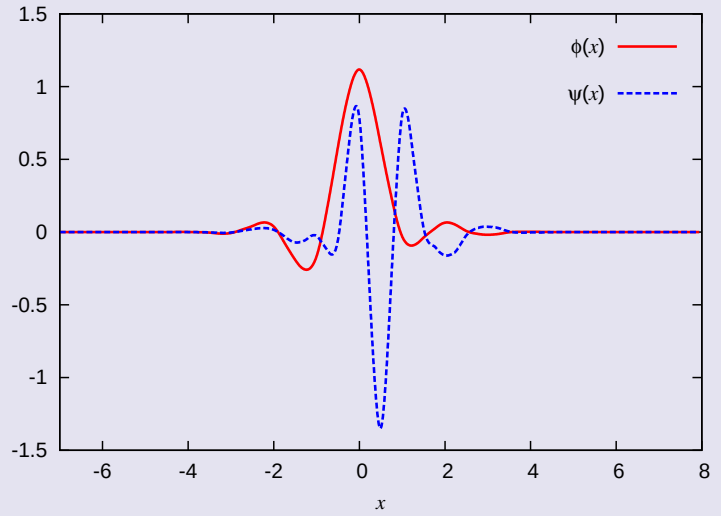
Wavelets

A basis with optimal properties for expanding localised information



- Localised in real space
- Smooth (localised in Fourier space)
- Orthogonal basis
- Multi-resolution basis
- Adaptive
- Systematic

Daubechies Wavelets



From early 80's

Applied in several domains

Interesting properties for DFT

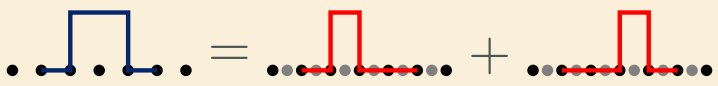
A brief description of wavelet theory

A Multi-Resolution real space basis

All functions w/ compact support, centered on grid points. In the wavelet theory we have two kind of basis functions.

Scaling Functions (SF)

The functions of low resolution level are a linear combination of high-resolution functions.



Wavelets (W)

Contain the DoF needed to complete the information lacking due to the coarseness of the resolution.

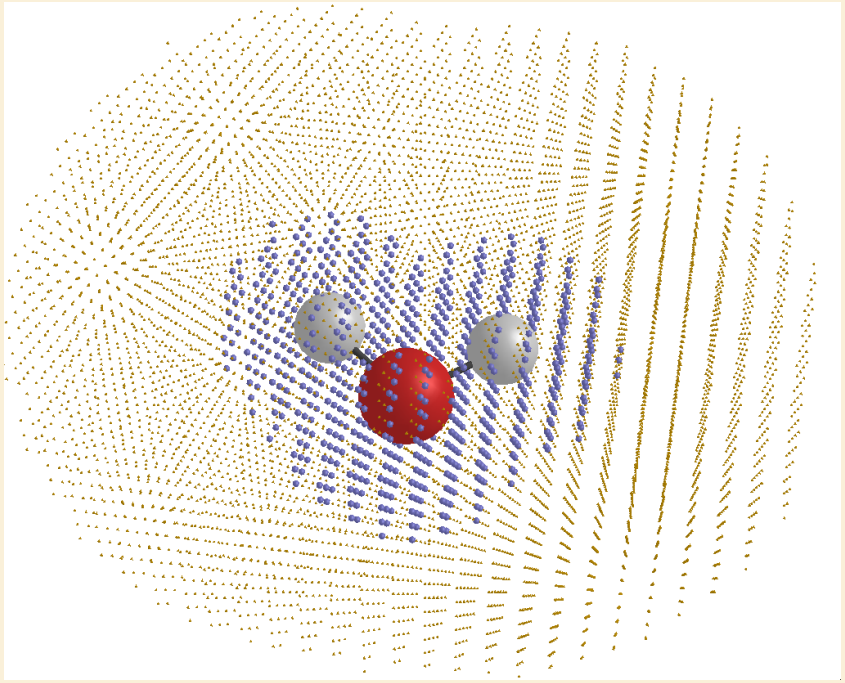
Increase the resolution without modifying grid space

SF + W = Degrees of Freedom of SF of higher resolution

Wavelet properties: adaptivity

Adaptivity

Resolution can be refined following the grid point.



The grid is divided in Low (1 DoF) and High (8 DoF) resolution points.
Points of different resolution belong to the same grid.
Empty regions must not be "filled" with basis functions.

Localization property,real space description

Optimal for big & inhomogeneous systems, highly flexible

Basis set features

Tensor product decomposition of the basis

The 3D basis is *separable* in 1D SF/ W.

$$\phi_{j_x,j_y,j_z}^{e_x,e_y,e_z}(x,y,z) = \phi_{j_x}^{e_x}(x)\phi_{j_y}^{e_y}(y)\phi_{j_z}^{e_z}(z)$$

(j_x, j_y, j_z) are the grid points, $\phi_j^{(0)}$ and $\phi_j^{(1)}$ the SF and the W.

Orthogonality, scaling relation

Daubechies wavelets are orthogonal and multi-resolution

$$\int \mathrm{d}x \phi_k(x)\phi_\ell(x) = \delta_{k\ell} \qquad \phi(x) = \frac{1}{\sqrt{2}} \sum_{j=-m}^{m} \mathrm{h}_j \phi(2x-j)$$

Hamiltonian-related quantities are calculated analytically

The accuracy is only limited by the basis set ($O(h_{\mathrm{grid}}^{14})$)

Operator expressions, pseudopotentials

Exact evaluation of kinetic energy

Expressed analytically by a convolution:

$$f(x) = \sum_\ell c_\ell \phi_\ell(x), \qquad \nabla^2 f(x) = \sum_\ell \tilde{c}_\ell \phi_\ell(x),$$

$$\tilde{c}_\ell = \sum_j c_j a_{\ell-j}, \qquad a_\ell \equiv \int \phi_0(x) \partial_x^2 \phi_\ell(x),$$

From N^3 to $3N$ calculations for separable objects

We save computational time when performing scalar products with separable functions (e.g. gaussians).

$$\int \mathrm{d}\mathbf{r}\, \psi(\mathbf{r}) e^{-\frac{1}{2}\left(\frac{r}{r_a}\right)^2} = \sum_{ijk} c_{ijk} d_i d_j d_k, \qquad d_i = \int \phi_i(x) e^{-\frac{1}{2}\left(\frac{x}{r_a}\right)^2}$$

d_i coefficients can be calculated at machine precision.

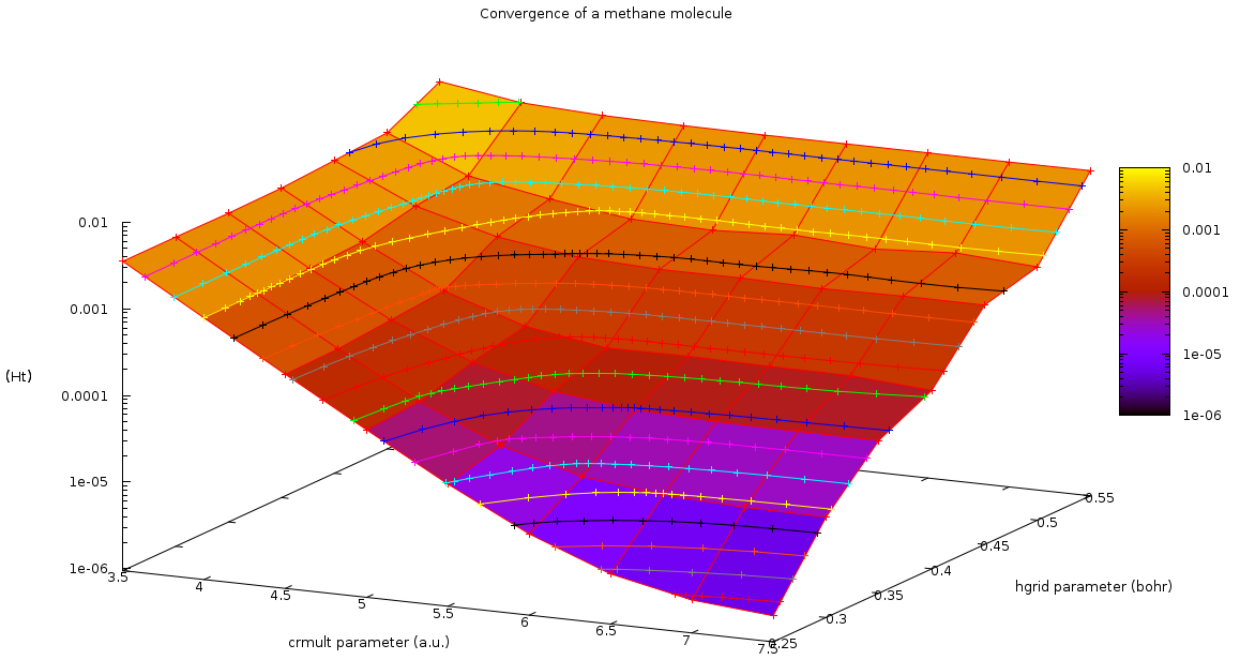
GTH-HGH: analytic and separable pseudopotentials

Systematic basis set

The absolute accuracy of the calculation is directly proportional to the number of the basis functions

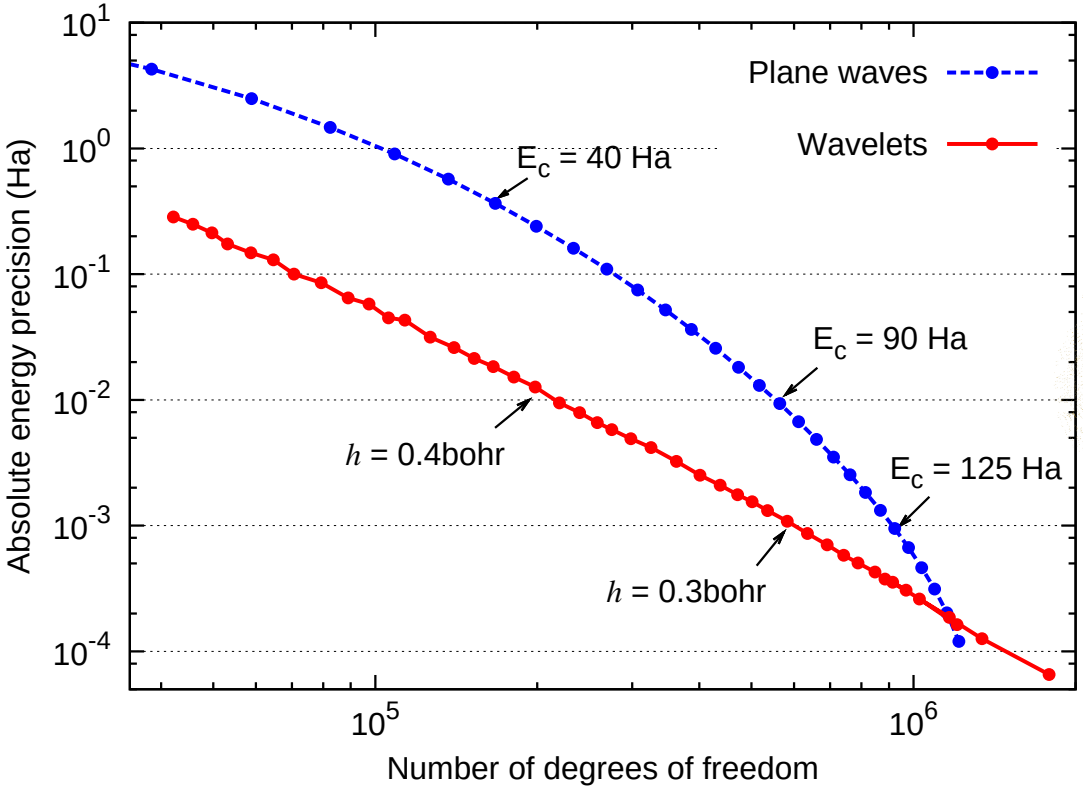
Two parameters for tuning the basis

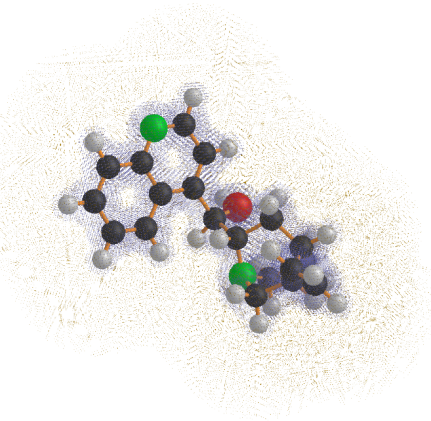
- The grid spacing `hgrid`
- The extension of the Low resolution points `crmult`



Optimal for inhomogeneous systems

Test case: cinchonidine molecule (44 atoms)





Enables a systematic approach for molecules

Considerably faster than Plane Waves codes.
the above run :10 (5) times faster than ABINIT (CPMD)
Charged systems can be treated *explicitly* with the same time

All-electron precision with HGH PSP

Non-linear core correction

J. Chem. Phys. **138**, 104109 (2013)

- Simple analytic form (a single gaussian as ρ_c)
- Same hardness as HGH
  ☛ a systematic localized basis is fundamental

Precision considerably improved

G2-1 test set (Atomization energies)

kcal/mol	MAD	RMSD	MSD	maxAD	minAD
Old HGH	6.85	9.13	-6.76	23.94	0.10
NLCC-HGH	**0.51**	**0.63**	**0.16**	**1.50**	**0.03**
PAW Paier	0.46	0.56	-0.43	1.13	0.01
Δ AE (geopt)	0.29	0.70	-0.29	4.21	0.00

AE precision for quantities in different environments

Bond lengths, Pressure (Bulk systems), Dispersion-corrected interaction energies, ...

Local orbitals and linear scaling

KS orbitals

Linear combinations of support functions $\phi_\alpha(\mathbf{r})$:

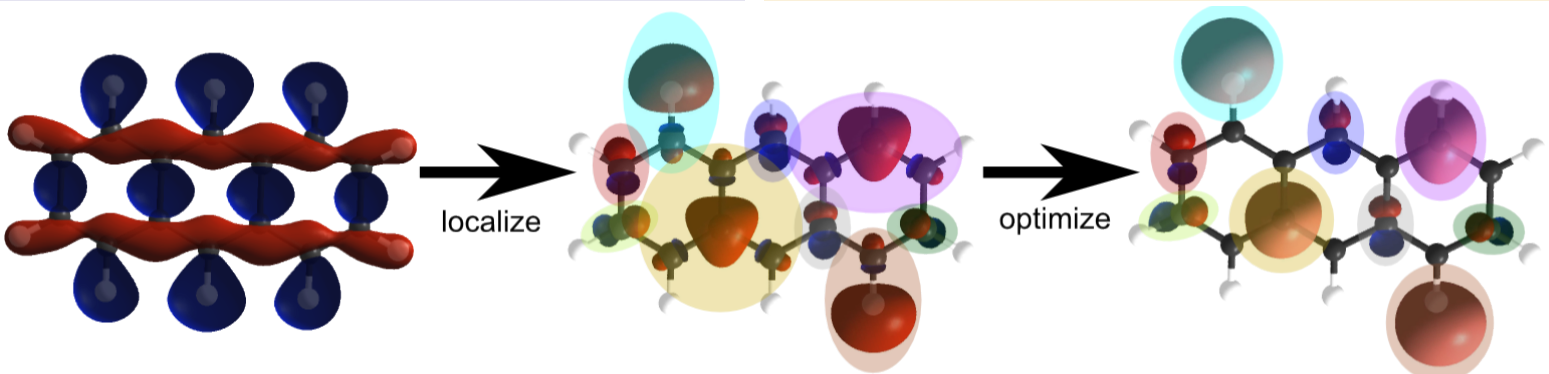
$$\Psi_i(\mathbf{r}) = \sum_\alpha c_i^\alpha \phi_\alpha(\mathbf{r})$$

- localized around atoms
- expanded in wavelets
- optimized in-situ

Density Matrix

Defined via the kernel $K^{\alpha\beta}$ in the $\phi_\alpha(\mathbf{r})$ basis:

$$\rho(\mathbf{r},\mathbf{r}') = \sum_i f_i \Psi_i(\mathbf{r})\Psi_i(\mathbf{r}')$$
$$= \sum_{\alpha,\beta} \phi_\alpha(\mathbf{r}) K^{\alpha\beta} \phi_\beta(\mathbf{r}')$$



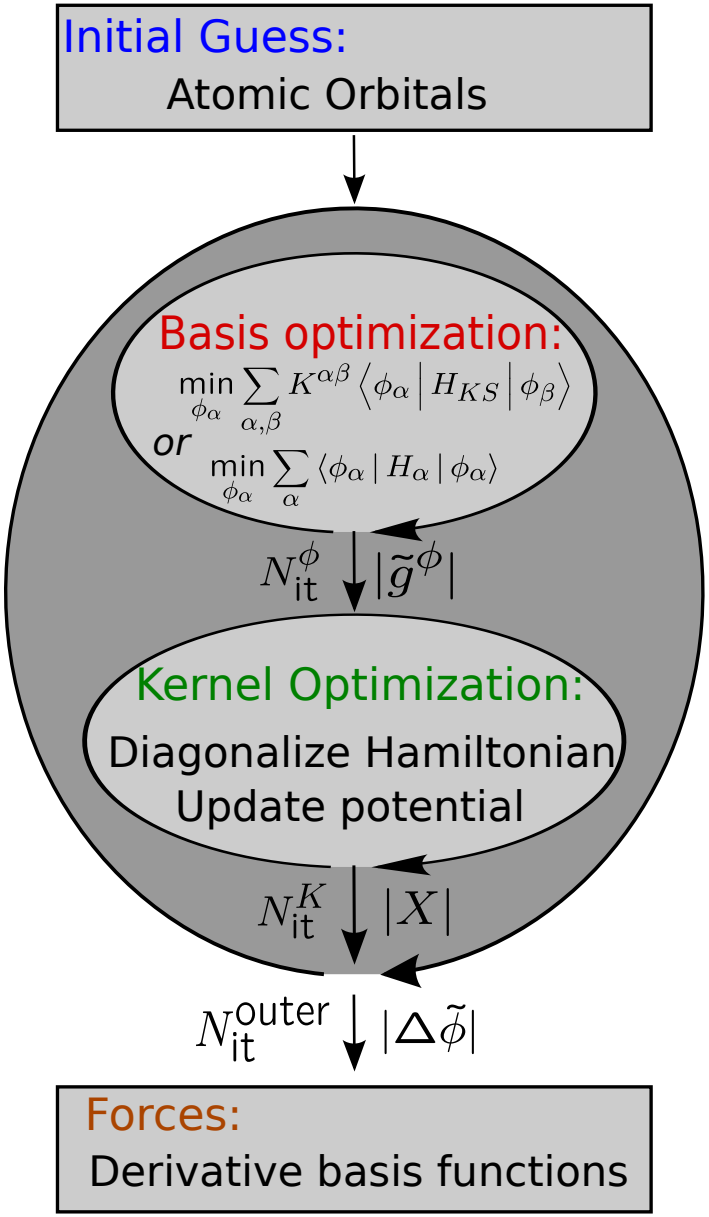
extended Kohn-Sham orbitals
cubic scaling, high accuracy

localized support functions (LCAO)
linear scaling, low accuracy

localized adaptive support functions
linear scaling, high accuracy

Localization $\rightarrow$ sparse matrices $\rightarrow O(N)$

Overall scheme



Two-step optimization scheme

- The ϕ_α minimize the "trace" of a confining KS hamiltonian
- Coefficients minimize KS energy

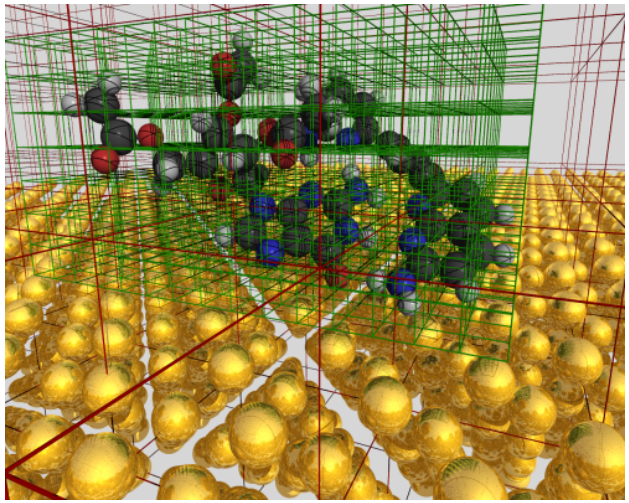
High quality results

- Good precision
- No need of Pulay forces due to basis completeness!

A flexible formalism

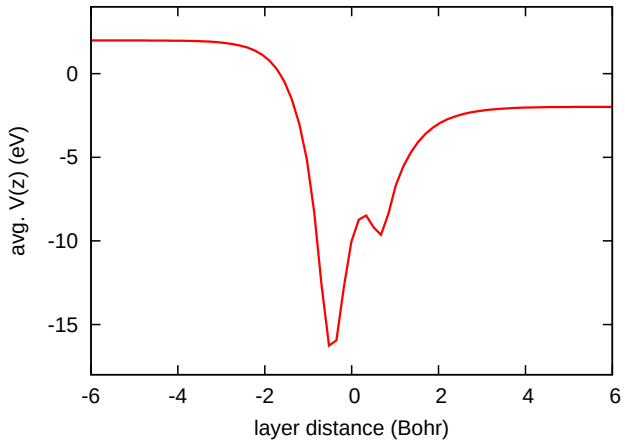
Flexible Boundary Conditions

- Isolated (free) BC
- Surfaces BC
- Periodic (3D) BC
- Wires BC



Systematic approach

Only relevant degrees of freedom are taken into account
Boundary conditions can be implemented explicitly



E.g.: Surfaces BC

2D Periodic + 1D isolated
Optimal to treat dipolar systems
without corrections