

# **Advanced Materials Modeling**

## **Lecture of 21 april**

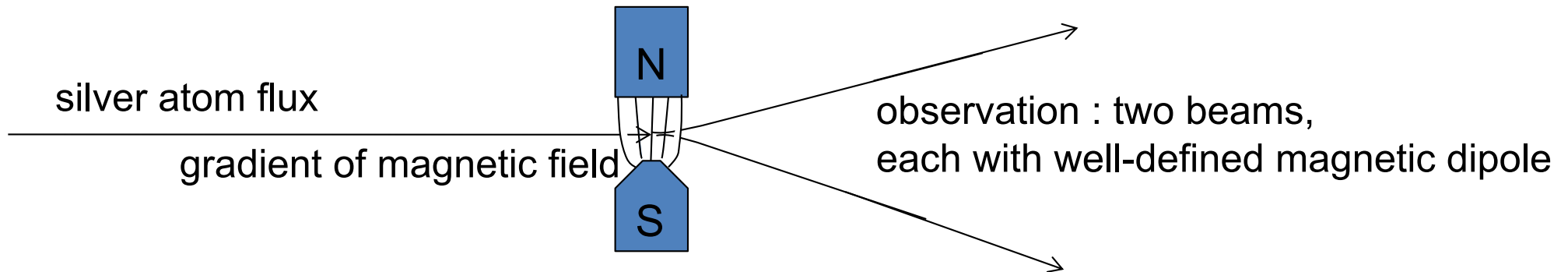
### **Part A**

- + Spin, magnetism, spin-orbit coupling, spin-density functional theory
- + Forces, and brief introduction to density functional perturbation theory
- + Iterative techniques
- + Troubleshooting

# Spin, magnetism, spin-orbit coupling, spin-density functional theory

# Intrinsic angular momentum

**Spin** Notion proposed by Uehlenbeck et Goudsmit (1925)  
to explain Stern and Gerlach experiment



Silver  $\text{Ag} = (\text{Kr}) + 4d^{10} 5s^1$

Intrinsic angular momentum



Behaves like an orbital moment, although with only two possible  
values of the  $z$  projection :  $m_s \hbar$   $m_s = \pm 1/2$

No intrinsic angular momentum in classical mechanics,  
no correspondance principle ...

# Reminder : orbital momentum

- Vector property  $\vec{\ell} = \vec{r} \times \vec{p} \rightarrow \hat{\ell} = -i\hbar(\hat{r} \times \hat{\nabla})$   
$$\hat{\ell}_z = -i\hbar \left( \hat{x} \frac{\partial}{\partial y} - \hat{y} \frac{\partial}{\partial x} \right)$$

Quantum eigenfunctions and eigenvalues

$$\hat{\ell}_z \varphi = \ell_z \varphi \qquad -i\hbar \left( x \frac{\partial \varphi}{\partial y} - y \frac{\partial \varphi}{\partial x} \right) = \ell_z \varphi$$

In spherical coordinates

$$\varphi(r, \theta, \Phi) = f(r, \theta) e^{i\ell_z \Phi / \hbar}$$

$$\ell_z = m \cdot \hbar \qquad m = 0, \pm 1, \pm 2, \dots \qquad \text{Discrete spectrum of } \hat{\ell}_z$$

# Representation of spin

Representation based on **vectors** (wavefunction) and **matrices** (operator) : intrinsic angular momentum cannot be represented using spatially-dependent wavefunction

Spinorial wavefunction: 2-component vector, called “spinor”

$$\Psi(\mathbf{r}) = \begin{pmatrix} \psi_{\uparrow}(\mathbf{r}) \\ \psi_{\downarrow}(\mathbf{r}) \end{pmatrix} \text{ or } \psi(\mathbf{r}, \sigma)$$

↪  $\sigma$  can take two “values”:  $\uparrow$  or  $\downarrow$

Other notation : define  $|\alpha\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$  et  $|\beta\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

$$\alpha(\uparrow) = 1 \quad \alpha(\downarrow) = 0 \quad \beta(\uparrow) = 0 \quad \beta(\downarrow) = 1$$

$$\psi(\mathbf{r}, \sigma) = \psi_{\uparrow}(\mathbf{r})\alpha(\sigma) + \psi_{\downarrow}(\mathbf{r})\beta(\sigma)$$

# Norm / scalar product

## Norm

$$N^2 = \langle \Psi | \Psi \rangle = \int \left( \psi_{\uparrow}(\mathbf{r}) \quad \psi_{\downarrow}(\mathbf{r}) \right)^* \begin{pmatrix} \psi_{\uparrow}(\mathbf{r}) \\ \psi_{\downarrow}(\mathbf{r}) \end{pmatrix} d\mathbf{r} = \int |\psi_{\uparrow}(\mathbf{r})|^2 + |\psi_{\downarrow}(\mathbf{r})|^2 d\mathbf{r}$$

## Scalar product

$$\begin{aligned} \langle \Phi | \Psi \rangle &= \int \left( \varphi_{\uparrow}(\mathbf{r}) \quad \varphi_{\downarrow}(\mathbf{r}) \right)^* \begin{pmatrix} \psi_{\uparrow}(\mathbf{r}) \\ \psi_{\downarrow}(\mathbf{r}) \end{pmatrix} d\mathbf{r} \\ &= \int \varphi_{\uparrow}^*(\mathbf{r}) \psi_{\uparrow}(\mathbf{r}) + \varphi_{\downarrow}^*(\mathbf{r}) \psi_{\downarrow}(\mathbf{r}) d\mathbf{r} \end{aligned}$$

# Pure spin-up or -down wavefunctions

Most often : pure spin up wavefunctions, or pure spin down wavefunctions

Explicitly  $\Psi(\mathbf{r}) = \begin{pmatrix} \psi_{\uparrow}(\mathbf{r}) \\ 0 \end{pmatrix}$  or  $\Psi(\mathbf{r}) = \begin{pmatrix} 0 \\ \psi_{\downarrow}(\mathbf{r}) \end{pmatrix}$

Such pure spin-up or spin-down wavefunctions are orthogonal to each others.  
Correspond to different quantum states.

Simply using spin-up or spin-down wavefunctions often referred to “scalar wavefunctions”, at variance with “spinorial wavefunctions” or “spinors”

# If not pure-up or -down wavefunctions ?

Computation of electronic density, and also spin-density.

Expectation value of electronic density at a point :

$$n(\mathbf{r}) = |\psi_{\uparrow}(\mathbf{r})|^2 + |\psi_{\downarrow}(\mathbf{r})|^2$$

Expectation value of the spin along  $z$  (spin-polarization)

$$s_z(\mathbf{r}) = \left( |\psi_{\uparrow}(\mathbf{r})|^2 - |\psi_{\downarrow}(\mathbf{r})|^2 \right) \frac{\hbar}{2}$$

What about the  $x$ - or  $y$ - spin-polarization ?  
Or expectation along any other direction ?



# Pauli matrices : spin along $x, y, z + \dots$

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Warning. Potential notation conflict :  $\sigma$  is for  $\uparrow$  or  $\downarrow$  ,  $\sigma_x, \sigma_y, \sigma_z$  are Pauli matrices

Spin-polarization vector (expectation value):

$$s_x(\mathbf{r}) = \left( \psi_{\downarrow}^*(\mathbf{r})\psi_{\uparrow}(\mathbf{r}) + \psi_{\uparrow}^*(\mathbf{r})\psi_{\downarrow}(\mathbf{r}) \right) \frac{\hbar}{2} = \Psi^+(\mathbf{r})\sigma_x\Psi(\mathbf{r})\frac{\hbar}{2}$$
$$s_y(\mathbf{r}) = \left( i\psi_{\downarrow}^*(\mathbf{r})\psi_{\uparrow}(\mathbf{r}) - i\psi_{\uparrow}^*(\mathbf{r})\psi_{\downarrow}(\mathbf{r}) \right) \frac{\hbar}{2} = \Psi^+(\mathbf{r})\sigma_y\Psi(\mathbf{r})\frac{\hbar}{2}$$
$$s_z(\mathbf{r}) = \left( |\psi_{\uparrow}(\mathbf{r})|^2 - |\psi_{\downarrow}(\mathbf{r})|^2 \right) \frac{\hbar}{2} = \Psi^+(\mathbf{r})\sigma_z\Psi(\mathbf{r})\frac{\hbar}{2}$$

Spin-density matrix :  $n_{\alpha\beta}(\mathbf{r}) = \psi_{\alpha}^*(\mathbf{r})\psi_{\beta}(\mathbf{r})$

$$n(\mathbf{r}) = n_{\uparrow\uparrow}(\mathbf{r}) + n_{\downarrow\downarrow}(\mathbf{r}) \quad s_z(\mathbf{r}) = \left( n_{\uparrow\uparrow}(\mathbf{r}) - n_{\downarrow\downarrow}(\mathbf{r}) \right) \frac{\hbar}{2}$$

Notation, often :  $n_{\uparrow}(\mathbf{r}) \triangleq n_{\uparrow\uparrow}(\mathbf{r}) \quad n_{\downarrow}(\mathbf{r}) \triangleq n_{\downarrow\downarrow}(\mathbf{r})$

# Operators acting on spinors

$$\hat{\mathbb{A}} = \begin{pmatrix} \hat{A}_{\uparrow\uparrow} & \hat{A}_{\uparrow\downarrow} \\ \hat{A}_{\downarrow\uparrow} & \hat{A}_{\downarrow\downarrow} \end{pmatrix} \quad \text{where } \hat{A}_{\uparrow\uparrow}, \hat{A}_{\uparrow\downarrow}, \hat{A}_{\downarrow\uparrow}, \hat{A}_{\downarrow\downarrow} \text{ are operators acting in usual Hilbert space (=spatial part)}$$

$$\hat{\mathbb{A}}|\Psi\rangle = \begin{pmatrix} \hat{A}_{\uparrow\uparrow}|\psi_{\uparrow}\rangle + \hat{A}_{\uparrow\downarrow}|\psi_{\downarrow}\rangle \\ \hat{A}_{\downarrow\uparrow}|\psi_{\uparrow}\rangle + \hat{A}_{\downarrow\downarrow}|\psi_{\downarrow}\rangle \end{pmatrix}$$

Simple example (simple, because spatial part = unit operator) :

$$\hat{S}_z = \begin{pmatrix} \frac{\hbar}{2}\hat{1} & 0 \\ 0 & -\frac{\hbar}{2}\hat{1} \end{pmatrix} = \frac{\hbar}{2}\sigma_z\hat{1} \quad \hat{S}_z|\alpha\rangle = \frac{\hbar}{2}|\alpha\rangle \quad \hat{S}_z|\beta\rangle = -\frac{\hbar}{2}|\beta\rangle$$

Generalization for  $x$  and  $y$  or other directions using Pauli matrices

# Classical treatment of a magnetic field

Now, how does spin relates to magnetic dipole ?

Classical mechanics: for a particle rotating around origin, with mass  $M$  and charge  $q$  :

Angular momentum  $\mathbf{L} = M\mathbf{r} \times \mathbf{v}$

Magnetic momentum  $\boldsymbol{\mu} = \frac{q}{2}\mathbf{r} \times \mathbf{v}$

yields  $\boldsymbol{\mu} = \frac{q}{2M}\mathbf{L}$

Placed inside a magnetic field,  $\mathbf{B}$   
a magnetic dipole acquire an energy  $E_{\text{int}} = -\boldsymbol{\mu} \cdot \mathbf{B}$

# Quantum treatment of a magnetic field

$$\mu = \frac{q}{2M} \mathbf{L} \quad + \text{Correspondance principle} \quad \Rightarrow \quad \hat{\mu} = \frac{q}{2M} \hat{\mathbf{L}}$$

Suppose a magnetic field along  $z$ .

For an electron (charge  $q=-e$ , mass  $m_e$ ), observable values for the projection of magnetic moment along  $z$  are

$$\mu_z = -\frac{e}{2m_e} L_z = -\frac{e}{2m_e} m\hbar$$

where  $m$  = magnetic quantum number (integer)

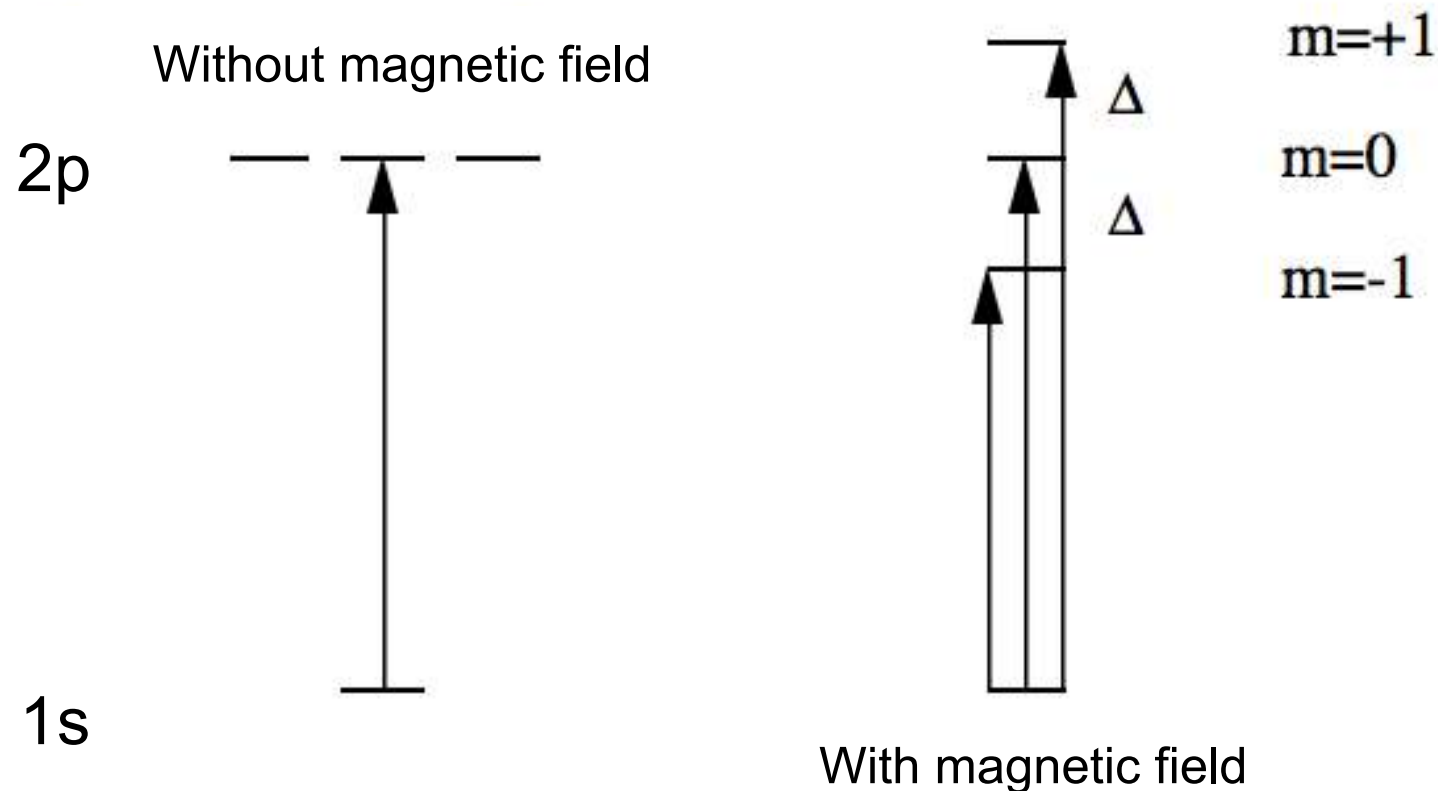
Definition of 'Bohr magneton'  $\mu_B = \frac{e\hbar}{2m_e}$

Energy change (first approximation)

$$E_{\text{int}} = -\mu_z \cdot B_z = \frac{e}{2m_e} m\hbar B_z = \frac{e\hbar}{2m_e} m B_z = \mu_B m B_z$$

# Zeeman effect : lifting the degeneracy

Atomic levels of same  $n$  and  $\ell$  , but different  $m$ , are not degenerate anymore in a magnetic field



# Intrinsic magnetic moment

Classical relation between orbital moment and magnetic moment :

$$\hat{\mu} = \frac{q}{2M} \hat{\mathbf{L}}$$

There exists also a spin magnetic moment, but the proportionality coefficient is not the same :

$$\hat{\mu}_s = g_e \frac{(-e)}{2m_e} \hat{\mathbf{S}}$$

g-factor :  $g_e \cong 2$

$$E_{\text{int}} = -\mu_z \cdot B_z = \frac{g_e e}{2m_e} m \hbar B_z = \frac{e \hbar}{2m_e} g_e m B_z = \mu_B g_e m B_z$$

If gradient of B,  
a force appears  
=> Stern-Gerlach exp !

# Spin-orbit (SOC) coupling (I)

Classically, magnetic dipoles interact with each others.

(magnetic dipole induces a magnetic field, felt by the other dipole)

Magnetic dipole from orbital movement interacts with spin magnetic dipole :

$$E_{soc} \propto \mathbf{L} \cdot \mathbf{S}$$

Atomic case : degeneracies are lifted.

Example :  $p$  levels  $m_\ell = -1, 0, +1$   $m_s = -\frac{1}{2}, +\frac{1}{2}$

$$\left\{ |p_{m=-1}, \uparrow\rangle, |p_{m=-1}, \downarrow\rangle, |p_{m=0}, \uparrow\rangle, |p_{m=0}, \downarrow\rangle, |p_{m=+1}, \uparrow\rangle, |p_{m=+1}, \downarrow\rangle \right\}$$

Due to SOC, 6-fold degeneracy becomes doublet+quadruplet

$$m_j \hbar = m_\ell \hbar + m_s \hbar$$

$$j = \frac{1}{2} \text{ with } m_j = -\frac{1}{2}, \frac{1}{2} \quad j = \frac{3}{2} \text{ with } m_j = -\frac{3}{2}, -\frac{1}{2}, \frac{1}{2}, \frac{3}{2}$$

$$p_{1/2} - p_{3/2} \text{ splitting}$$

# Spin-orbit interaction (II)

Theory : originates from relativistic corrections (Schrödinger to Dirac Eq.)

Mass-velocity + Darwin term

= modification of kinetic energy and potential

Spin-orbit interaction

= new type of term  $E_{soc} = -\left(\frac{e\hbar}{2m_e c}\right) \boldsymbol{\sigma} \cdot \left[ \nabla V \times \left( \frac{\mathbf{p}}{m_e c} \right) \right]$   $\boldsymbol{\sigma}$  = vector of Pauli matrices

Close to a nucleus,  $E_{soc} \propto \mathbf{L} \cdot \mathbf{S}$  with an energy modif proportional to  $Z^4$

Note : vanishes for s electronic states  $\ell = 0$

Often NOT taken into account in first-principles calculations, except for electronic properties. Needs spinorial wavefunctions, increase of CPU time > 4.

										18
										2
										He
										4.0026
										10
										Ne
										20.179
										18
										Ar
										39.948
										36
										Kr
										83.80
										54
										Xe
										131.29
										126.9045
										At
										(222)
										86
										Rn
										208.9804
										Bi
										(209)
										Po
										(210)
										At
										(210)
										Po
										(209)
										Bi
										208.9804
										Pb
										207.2
										Tl
										204.383
										Hg
										200.59
										Au
										196.9665
										Pt
										195.08
										78
										Pt
										106.42
										Pd
										46
										Ag
										107.868
										Cd
										112.41
										In
										114.82
										Sn
										118.69
										Sb
										121.75
										Te
										127.60
										I
										126.9045
										Br
										79.904
										Selenium
										78.96
										34
										Se
										32.06
										S
										16
										Cl
										35.453
										17
										Ar
										39.948
										18
										Ne
										20.179
										10
										He
										4.0026

Z=78 ... 83 →

$6p_{1/2} - 6p_{3/2}$  splitting in Bi is about 1.5 eV

Incomplete p-shell



# Exchange-correlation functionals

In principle, exact functional is still a functional of density only !

$$E_{xc}^{\text{DFT}}[n] = \int n(\mathbf{r}) \mathcal{E}_{xc}^{\text{DFT}}(\mathbf{r}; n) d\mathbf{r}$$

**However**, information on density + **spin** gives better approximate functionals.  
“Spin-density functional theory” (SDFT).

O. Gunnarsson and B.I. Lunqvist, *Phys. Rev. B* **13**, 4274 (1976)

If **spin-orbit coupling is ignored**, most materials have either **no spin-polarization** (nowhere in space), or their **spin-polarization is aligned with only one direction** (chosen as  $z$ ), everywhere in space.

The latter gives the **collinear case** :

$$E_{xc}^{\text{SDFT,coll}}[n_{\uparrow}, n_{\downarrow}] = \int n(\mathbf{r}) \mathcal{E}_{xc}^{\text{SDFT,coll}}(\mathbf{r}; n_{\uparrow}, n_{\downarrow}) d\mathbf{r}$$

If **spin-orbit coupling is taken into account**, either there is **no spin-polarization** (nowhere in space), or **spin-polarization is not constrained to be aligned**.

The latter gives the **non-collinear case** :

$$E_{xc}^{\text{SDFT,non-coll}}[n, \mathbf{s}] = \int n(\mathbf{r}_1) \mathcal{E}_{xc}^{\text{SDFT,non-coll}}(\mathbf{r}; n, \mathbf{s}) d\mathbf{r}$$

# Exchange-correlation potentials

Generalisation of Kohn-Sham equations :

**Collinear case** : pure spin-up or spin-down wavefunctions  $|\psi_i^\alpha\rangle$

$$\left(-\frac{1}{2}\nabla^2 + V_{\text{KS},\alpha}(\mathbf{r})\right)\psi_{i\alpha}(\mathbf{r}) = \varepsilon_{i\alpha}\psi_{i\alpha}(\mathbf{r})$$

$i$  and  $\alpha$  label the states  
 $\alpha = \uparrow$  or  $\downarrow$

$$V_{\text{KS},\alpha}(\mathbf{r}) = V_{\text{ext}}(\mathbf{r}) + \int \frac{n(\mathbf{r}_1)}{|\mathbf{r}_1 - \mathbf{r}|} d\mathbf{r}_1 + \frac{\delta E_{\text{xc}}^{\text{SDFT, coll}}[n_\uparrow, n_\downarrow]}{\delta n_\alpha(\mathbf{r})}$$

Spin-dependent XC potential

**Non-collinear case** : spinorial wavefunctions  $|\Psi_i\rangle$

$$\left(-\frac{1}{2}\nabla^2 + V_{\text{KS}}(\mathbf{r}) + \frac{\hbar}{2}\mathbf{B}_{\text{xc}}(\mathbf{r}) \cdot \boldsymbol{\sigma}\right)\Psi_i(\mathbf{r}) = \varepsilon_i\Psi_i(\mathbf{r})$$

$i$  labels the states

$$V_{\text{KS}}(\mathbf{r}) = V_{\text{ext}}(\mathbf{r}) + \int \frac{n(\mathbf{r}_1)}{|\mathbf{r}_1 - \mathbf{r}|} d\mathbf{r}_1 + \frac{\delta E_{\text{xc}}^{\text{SDFT, non-coll}}[n, \mathbf{s}]}{\delta n(\mathbf{r})}$$

$$B_{\text{xc},j}(\mathbf{r}) = \frac{\delta E_{\text{xc}}^{\text{SDFT, non-coll}}[n, \mathbf{s}]}{\delta s_j(\mathbf{r})}$$

$j = x, y, z$       XC magnetic field

# Approximate XC functionals

LDA, GGA, hybrids : can easily be generalized.

E.g. Local spin-density approximation for the **collinear case** :

$$E_{xc}^{\text{LSDA}}[n_{\uparrow}, n_{\downarrow}] = \int n(\mathbf{r}) \varepsilon_{xc}^{\text{LSDA}}(n_{\uparrow}(\mathbf{r}), n_{\downarrow}(\mathbf{r})) d\mathbf{r}$$

$\varepsilon_{xc}^{\text{LSDA}}(n_{\uparrow}(\mathbf{r}), n_{\downarrow}(\mathbf{r}))$  parameterized from homogeneous electron gas

Moreover, exchange energy in **collinear case** has simple general relationship with unpolarized exchange energy :

G.L. Oliver and J.P. Perdew, *Phys. Rev. A* **20**, 397 (1979)

$$E_x^{\text{SDFT, coll}}[n_{\uparrow}, n_{\downarrow}] = \frac{1}{2} \left( E_x^{\text{DFT}}[2n_{\uparrow}] + E_x^{\text{DFT}}[2n_{\downarrow}] \right)$$

Indeed, comes from separate contribution of spins to exchange energy:

$$E_x = -\frac{1}{2} \sum_{\sigma} \sum_{i,j}^{\text{occ}} \int d\mathbf{r} d\mathbf{r}' \frac{\psi_{j\sigma}^*(\mathbf{r}') \psi_{i\sigma}^*(\mathbf{r}') \psi_{j\sigma}(\mathbf{r}) \psi_{i\sigma}(\mathbf{r})}{|\mathbf{r} - \mathbf{r}'|}$$

For **non-collinear case**, neglect directional dependence of local spin-polarization => falls back to collinear case.

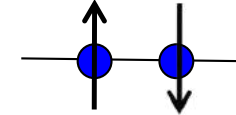
# Methodologies : spin+symmetry

In practice ... ?

Materials are treated within one of the following methodologies:

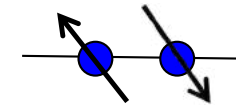
(1) No spin-polarization, scalar wavefunctions

(Many metals/covalent/ionic solids with light nuclei)



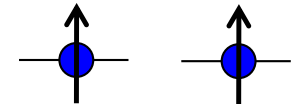
(2) No spin-polarization, spinorial wavefunctions (spin-orbit coupling)

(Many metals/covalent/ionic solids with heavier nuclei ;  
the spin-polarization vanishes because of non-broken  
time-reversal (TR) symmetry, compensation from TR symmetric states)



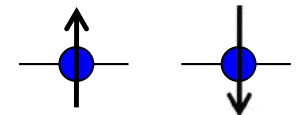
(3) Collinear spin, different spin-up and spin-down wavefunctions

(Ferromagnetic materials with light nuclei, finite systems with net spin,  
in particular open shell atoms, and paramagnetic molecules, if light nuclei)



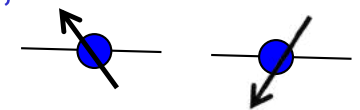
(4) Collinear spin, antiferromagnetic case

(Specific materials, often oxides – the spin-down density  
is space-symmetric to the spin-up density – spin-polarization does not vanish)



(5) Non-collinear spin-polarization

(Heavy nuclei with magnetization, need TR symm. spontaneously broken,  
also important for some properties like the magnetocrystalline field,  
or spin diffusion)



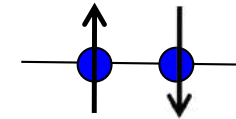
# Methodologies : spin+symmetry

In practice ... ?

ABINIT input variables `nspinor`, `nspden`, `nsppol` . Also `spinat(3,natom)`.

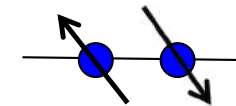
(1) No spin-polarization, scalar wavefunctions

`nspinor=1` `nspden=1` `nsppol=1`



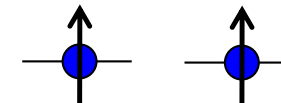
(2) No spin-polarization, spinorial wavefunctions (spin-orbit coupling)

`nspinor=2` `nspden=1` `nsppol=1`



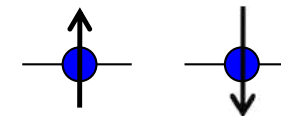
(3) Collinear spin, different spin-up and spin-down wavefunctions

`nspinor=1` `nspden=2` `nsppol=2`



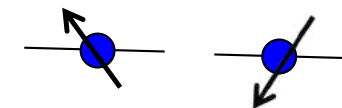
(4) Collinear spin, antiferromagnetic case

`nspinor=1` `nspden=2` `nsppol=1`

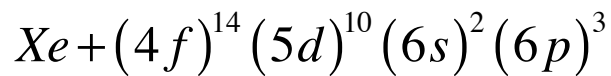


(5) Non-collinear spin-polarization

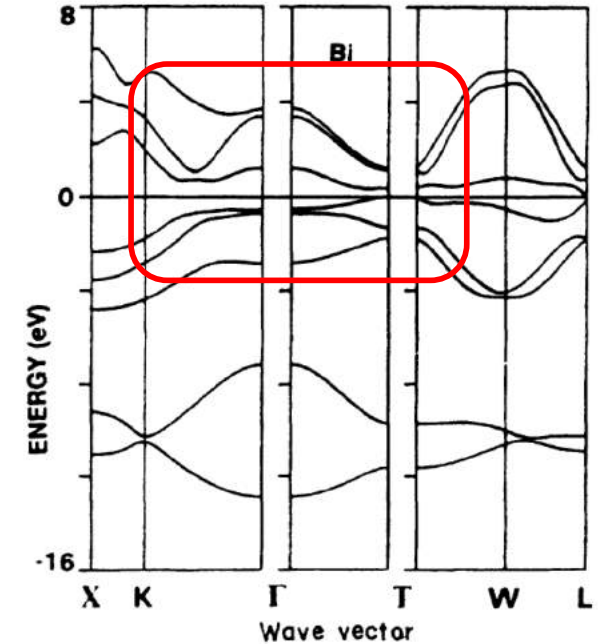
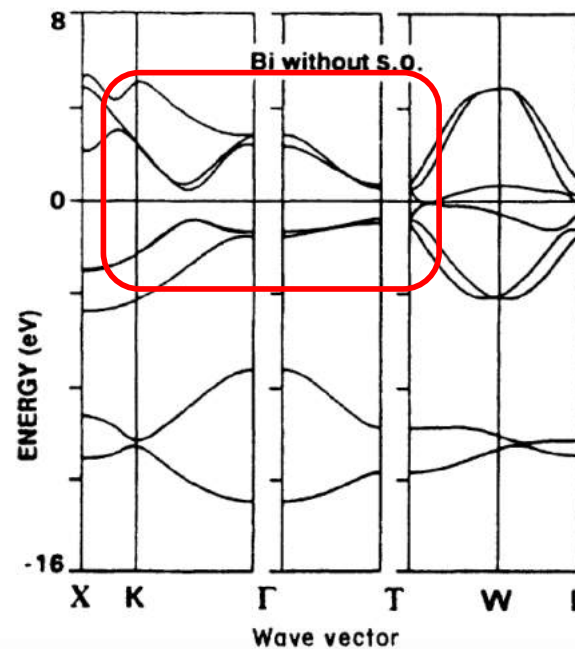
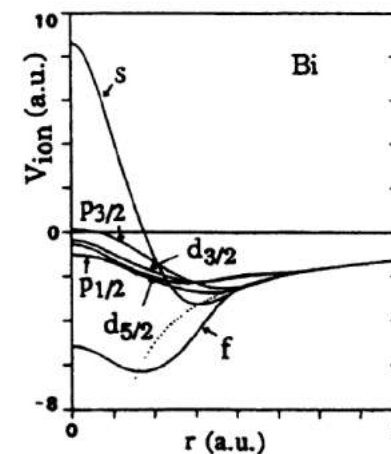
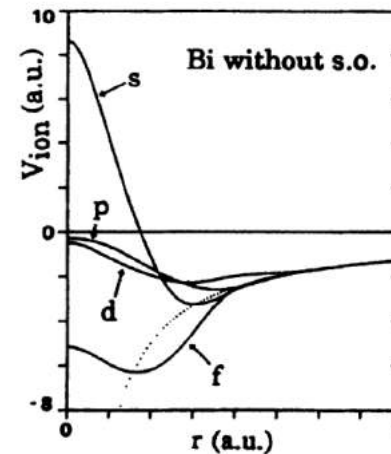
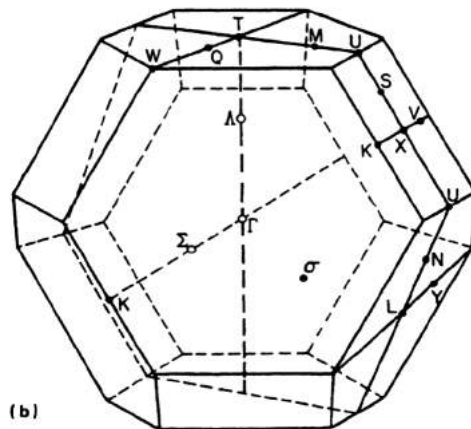
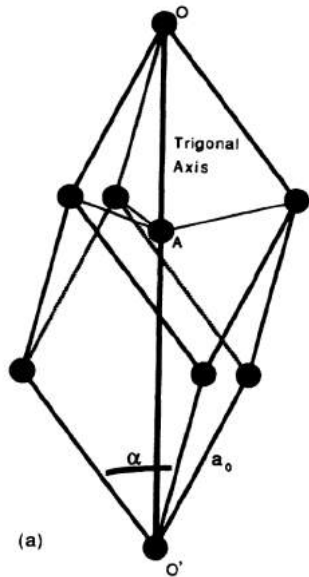
`nspinor=2` `nspden=4` `nsppol=1`



# Example : electronic structure of Bismuth



$6p_{1/2}$ - $6p_{3/2}$  splitting is about 1.5 eV



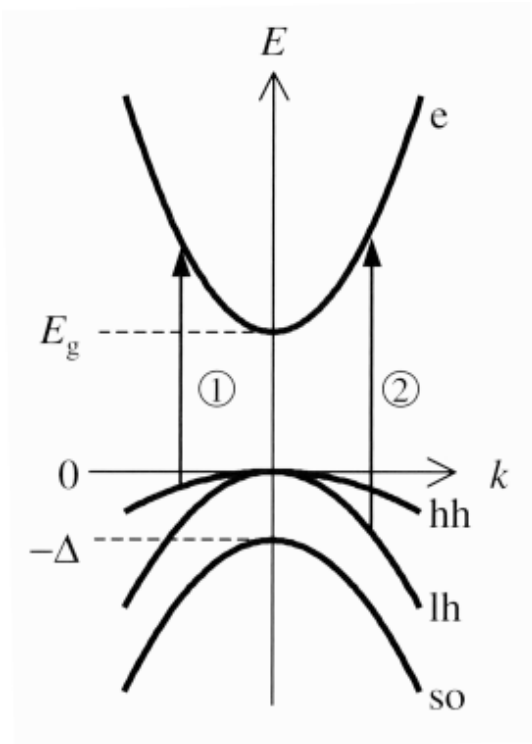
# Electronic structure of III-V & II-VI compounds

III-V and II-VI compounds : important for **opto-electronic applications**  
(light-emitting diodes, IR and visible photodetectors, solid state lasers)

Crystalline structure : tetrahedrally coordinated

(zinc-blende - AlP, AlAs, AlSb, BN, BP, CdTe, GaAs, HgSe, HgTe, InAs, InP, InSb, ZnS ;  
or wurtzite - AlN, GaN, InN, CdS, CdSe, CdTe, ZnO, ZnS )

ZB Band structure in the vicinity of the  $\Gamma$  point (zone center) : Kane model



## Kane model

Band structure of several III-V and II-VI compounds near  $k=0$ .

$E=0$  : top of the valence band

$E = E_g$  : bottom of the conduction band.

4 bands :

heavy hole ( $hh$ ) band, light hole ( $lh$ ) band, spin-orbit hole ( $so$ ) band, and electron ( $e$ ) band.

Two optical transitions are indicated.

Transitions can also take place between spin-orbit hole band and conduction band, but are not shown for the sake of clarity.

$\Delta$  ranges from  $<0.1$  eV to  $0.33$  eV for GaAs and  $0.8$  eV for InSb

# Role of XC functional : the case of Iron

Total energy as a function of volume per atom for Iron.

Dashed lines : Spin-LDA

Full lines : GGA

P = paramagnetic (no spin-polarization)

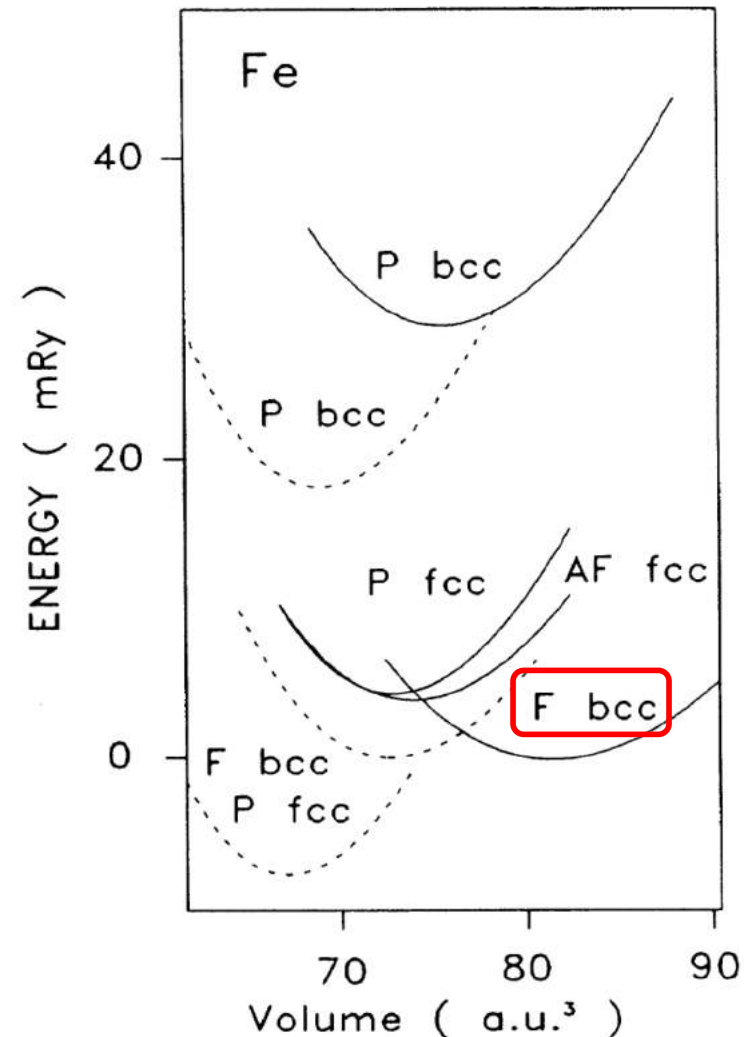
F = ferromagnetic

AF = antiferromagnetic

FCC= face-centered cubic

BCC= body-centered cubic

The GGA gets the correct phase :  
Experimentally Fe is ferromagnetic BCC.



T.C. Leung, C.T. Chan and B.N. Harmon, *Phys. Rev. B* **44**, 2923 (1991)



# Time-reversal operator in the scalar case

Many-body electronic Hamiltonian : invariant upon time-reversal symmetry.

$[\hat{H}_{el}, \hat{T}] = 0$  with  $\hat{T}$  time-reversal operator = simply taking the complex conjugate

(Think to time-dependent Schrödinger equation  $\hat{H}|\Psi_i(t)\rangle = i\hbar \frac{\partial}{\partial t} |\Psi_i\rangle$  )

$$\hat{T}^2 = \hat{1} \quad \hat{T}^{-1} = \hat{T} \quad [\hat{H}_{el}, \hat{T}] = 0 \Rightarrow \hat{T}\hat{H}_{el}\hat{T} = \hat{T}\hat{T}\hat{H}_{el} = \hat{H}_{el}$$

Upon time-reversal:  $\hat{\mathbf{r}}$  invariant,  $\hat{\mathbf{r}}\hat{T} = \hat{T}\hat{\mathbf{r}}$ , while  $\hat{\mathbf{p}}$  is reversed  $\hat{\mathbf{p}}\hat{T} = -\hat{T}\hat{\mathbf{p}}$

If Kohn-Sham Hamiltonian is invariant upon time-reversal symmetry, and if

$$\hat{H}_{KS}|\Psi_i\rangle = \varepsilon_i|\Psi_i\rangle \quad \text{then} \quad \hat{T}\hat{H}_{KS}\hat{T}|\Psi_i\rangle = \varepsilon_i|\Psi_i\rangle$$

$$\hat{H}_{KS}(\hat{T}|\Psi_i\rangle) = \varepsilon_i(\hat{T}|\Psi_i\rangle)$$

which means  $\hat{T}|\Psi_i\rangle$  also eigenfunction, with same eigenenergy than  $|\Psi_i\rangle$ .

Scalar wavefunctions in periodic boundary conditions and planewave basis

$$\psi_{\mathbf{k}}(\mathbf{r}) = (N\Omega_0)^{-1/2} \sum_{\mathbf{G}} u_{\mathbf{k}}(\mathbf{G}) e^{i(\mathbf{k}+\mathbf{G})\mathbf{r}}$$

Then there exists  $n'$  such that

$$\hat{T}\psi_{n\mathbf{k}}(\mathbf{r}) = \psi_{n\mathbf{k}}^*(\mathbf{r}) = \psi_{n'-\mathbf{k}}(\mathbf{r})$$

$$\varepsilon_{n\mathbf{k}}(\mathbf{r}) = \varepsilon_{n'-\mathbf{k}}(\mathbf{r})$$

# Time-reversal and spin

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

$\sigma_y$  not invariant upon complex conjugation. Problem with time-reversal symmetry ?

Invariance under complex conjugation AND exchange of spin-up and spin-down.  
Time-reversal symmetry exchange spin up and spin down (=angular momentum).

$$\hat{H}_{soc} = -\left(\frac{e\hbar}{2m_e c}\right) \boldsymbol{\sigma} \cdot \left[ \nabla V \times \left( \frac{\hat{\mathbf{p}}}{m_e c} \right) \right] \quad \text{moreover, } \hat{\mathbf{p}} \text{ changes sign}$$

If Kohn-Sham Hamiltonian including SOC is invariant upon time-reversal symmetry, and if

$$\hat{H}_{KS} \begin{pmatrix} \psi_{\uparrow n\mathbf{k}}(\mathbf{r}) \\ \psi_{\downarrow n\mathbf{k}}(\mathbf{r}) \end{pmatrix} = \epsilon_{n\mathbf{k}} \begin{pmatrix} \psi_{\uparrow n\mathbf{k}}(\mathbf{r}) \\ \psi_{\downarrow n\mathbf{k}}(\mathbf{r}) \end{pmatrix}$$

Then there exists  $n'$  such that  $\epsilon_{n\mathbf{k}}(\mathbf{r}) = \epsilon_{n'-\mathbf{k}}(\mathbf{r})$

and  $\begin{pmatrix} \psi_{\downarrow n\mathbf{k}}^*(\mathbf{r}) \\ -\psi_{\uparrow n\mathbf{k}}^*(\mathbf{r}) \end{pmatrix} = \begin{pmatrix} \psi_{\uparrow n'-\mathbf{k}}(\mathbf{r}) \\ \psi_{\downarrow n'-\mathbf{k}}(\mathbf{r}) \end{pmatrix}$  is the corresponding eigenfunction.  
= Kramers degeneracy

# Time-reversal symmetry forbids spin polarization

Compute expectation value of spin polarization from a pair of degenerate Kramers eigenfunctions :

$$\begin{pmatrix} \psi_{\downarrow n\mathbf{k}}^*(\mathbf{r}) \\ -\psi_{\uparrow n\mathbf{k}}^*(\mathbf{r}) \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \psi_{\uparrow n\mathbf{k}}(\mathbf{r}) \\ \psi_{\downarrow n\mathbf{k}}(\mathbf{r}) \end{pmatrix}$$

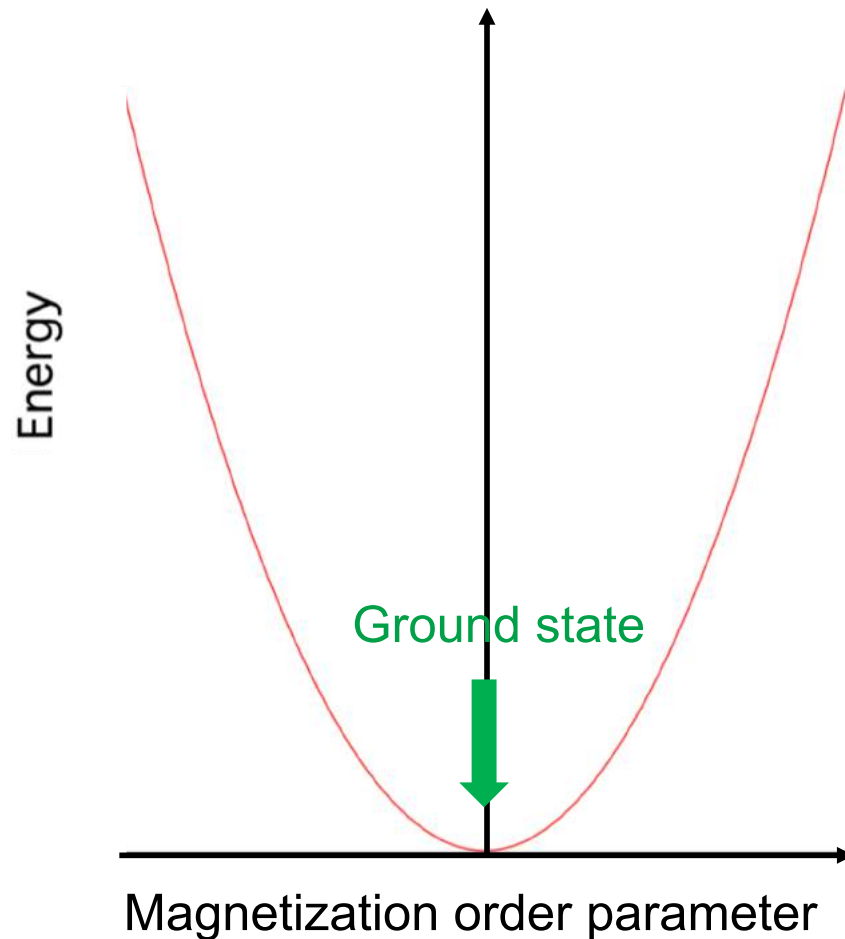
Reminder  $s_j(\mathbf{r}) = \Psi^\dagger(\mathbf{r}) \sigma_j \Psi(\mathbf{r}) \frac{\hbar}{2}$

This gives 
$$s_x(\mathbf{r}) = \left( -\psi_{\uparrow n\mathbf{k}}(\mathbf{r}) \psi_{\downarrow n\mathbf{k}}^*(\mathbf{r}) - \psi_{\downarrow n\mathbf{k}}(\mathbf{r}) \psi_{\uparrow n\mathbf{k}}^*(\mathbf{r}) \right) \frac{\hbar}{2} + \left( \psi_{\uparrow n\mathbf{k}}^*(\mathbf{r}) \psi_{\downarrow n\mathbf{k}}(\mathbf{r}) + \psi_{\downarrow n\mathbf{k}}^*(\mathbf{r}) \psi_{\uparrow n\mathbf{k}}(\mathbf{r}) \right) \frac{\hbar}{2} = 0$$

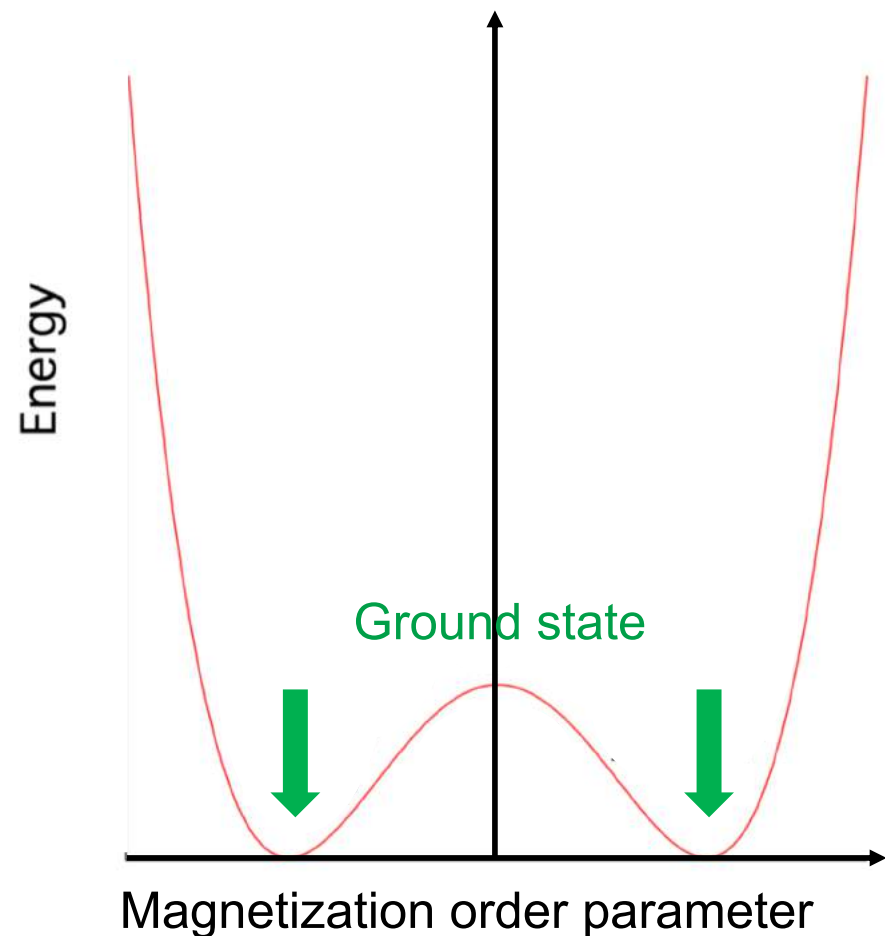
Similarly :  $s_y(\mathbf{r}) = 0$  and  $s_z(\mathbf{r}) = 0$

So, how can a material be magnetic ?

# Spontaneous breaking of time-reversal symmetry



Spin-polarization vanishes,  
Time-reversal symmetry not broken  
Non-magnetic, possibly spin texture

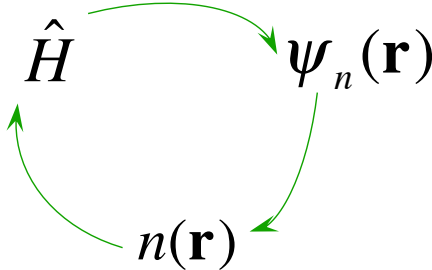


Spin-polarization does not vanish  
Time-reversal symmetry broken  
Magnetic materials

# Computing the forces

# Basic equations in DFT

Solve self-consistently the Kohn-Sham equation

$$\left\{ \begin{array}{l} \hat{H} |\psi_n\rangle = \epsilon_n |\psi_n\rangle \\ \hat{H} = \hat{T} + \hat{V} + \hat{V}_{Hxc}[n] \\ n(\mathbf{r}) = \sum_n^{occ} \psi_n^*(\mathbf{r}) \psi_n(\mathbf{r}) \end{array} \right.$$


$$\delta_{mn} = \langle \psi_m | \psi_n \rangle \text{ for } m, n \in \text{occupied set}$$

or minimize

$$E_{el} \{ \psi \} = \sum_n^{occ} \langle \psi_n | \hat{T} + \hat{V} | \psi_n \rangle + E_{Hxc}[n]$$

with

$$\hat{V}(\mathbf{r}) = \sum_{\kappa} -\frac{Z_{\kappa}}{|\mathbf{r} - \mathbf{R}_{\kappa}|}$$

Exchange-correlation functional might be LDA, GGA (e.g. PBE, PBESol), or hybrids, van der waals, etc

# Computing the forces (I)

Born - Oppenheimer approx.  $\Rightarrow$  find electronic ground state in potential created by nuclei.

A starting configuration of nuclei  $\{\mathbf{R}_\mathbf{k}\}$  is usually NOT in equilibrium geometry.

$$F_{\mathbf{k},\alpha} = - \left. \frac{\partial E}{\partial R_{\mathbf{k},\alpha}} \right|_{\{\mathbf{R}_\mathbf{k}\}} \quad (\text{principle of virtual works})$$

Forces are first derivatives of total energy.  
Can be computed by finite differences.

Better approach : compute the response to a perturbation

$\Rightarrow$  What is the energy change ?

$$\{\mathbf{R}_{\mathbf{k},\alpha}\} \rightarrow \{\mathbf{R}_{\mathbf{k},\alpha} + \lambda \delta \mathbf{R}_{\mathbf{k},\alpha}\}$$

*Small parameter*

# Computing the forces (II)

To simplify, let's compute the derivative of an electronic eigenvalue

Perturbation theory : Hellmann - Feynman theorem

$$\hat{H}(\lambda) \quad \varepsilon_n(\lambda) \quad \psi_n(\lambda) \text{ normalized} \quad \hat{H}|\psi_n\rangle = \varepsilon_n|\psi_n\rangle \quad \text{for all } \lambda$$

$$\begin{aligned} \frac{d\varepsilon_n}{d\lambda} &= \frac{d}{d\lambda} \left( \langle \psi_n | \hat{H} | \psi_n \rangle \right) \\ &= \left\langle \frac{d\psi_n}{d\lambda} | \hat{H} | \psi_n \right\rangle + \left\langle \psi_n | \frac{d\hat{H}}{d\lambda} | \psi_n \right\rangle + \left\langle \psi_n | \hat{H} | \frac{d\psi_n}{d\lambda} \right\rangle \\ &= \left\langle \frac{d\psi_n}{d\lambda} | \varepsilon_n | \psi_n \right\rangle + \left\langle \psi_n | \frac{d\hat{H}}{d\lambda} | \psi_n \right\rangle + \left\langle \psi_n | \varepsilon_n | \frac{d\psi_n}{d\lambda} \right\rangle \quad = \left\langle \psi_n^{(0)} | \frac{d\hat{H}}{d\lambda} | \psi_n^{(0)} \right\rangle \end{aligned}$$

$$\text{Indeed} \quad \left\langle \frac{d\psi_n}{d\lambda} | \psi_n \right\rangle + \left\langle \psi_n | \frac{d\psi_n}{d\lambda} \right\rangle = \frac{d}{d\lambda} \langle \psi_n | \psi_n \rangle = \frac{d}{d\lambda} (1) = 0$$

$$\frac{d\psi_n}{d\lambda} \quad \text{not needed !}$$



# Computing the forces (III)

Application to the derivative of  
with respect to an atomic displacement :

$$\hat{H} = \hat{T} + \hat{V}_{ext}\{\hat{\mathbf{R}}\} \Rightarrow \frac{\partial \hat{H}}{\partial \mathbf{R}_{\kappa,\alpha}} = \frac{\partial \hat{V}_{ext}}{\partial \mathbf{R}_{\kappa,\alpha}}$$

$$\frac{\partial \epsilon_n}{\partial \mathbf{R}_{\kappa,\alpha}} = \left\langle \psi_n \left| \frac{\partial \hat{H}}{\partial \mathbf{R}_{\kappa,\alpha}} \right| \psi_n \right\rangle = \int \frac{\partial \hat{V}_{ext}(\mathbf{r})}{\partial \mathbf{R}_{\kappa,\alpha}} n(\mathbf{r}) d\mathbf{r}$$

# Computing the forces (IV)

Generalisation to density functional theory

**Reminder :**  $E[\psi] = \sum_n \langle \psi_n | \hat{T} | \psi_n \rangle + \int n(\mathbf{r}) V_{\text{ext}}(\mathbf{r}) d\mathbf{r} + E_{\text{Hxc}}[n]$

If change of atomic positions ...

$$V_{\text{ext}}(\mathbf{r}) = \sum_k -\frac{Z_k}{|\mathbf{r} - \mathbf{R}_k|} \quad (\text{can be generalized to pseudopotential case})$$

$$\frac{\partial V_{\text{ext}}(\mathbf{r})}{\partial R_{k,\alpha}} = +\frac{Z_k}{|\mathbf{r} - \mathbf{R}_k|^2} \cdot \frac{\partial |\mathbf{r} - \mathbf{R}_k|}{\partial R_{k,\alpha}} = -\frac{Z_k}{|\mathbf{r} - \mathbf{R}_k|^3} \cdot (\mathbf{r} - \mathbf{R}_k)_\alpha$$

$$\boxed{\frac{\partial E}{\partial R_{k,\alpha}} = \int n(\mathbf{r}) \frac{\partial V_{\text{ext}}(\mathbf{r})}{\partial R_{k,\alpha}} d\mathbf{r} = - \int \frac{n(\mathbf{r})}{|\mathbf{r} - \mathbf{R}_k|^3} \cdot (\mathbf{r} - \mathbf{R}_k)_\alpha d\mathbf{r}}$$

**Forces can be computed directly from the density !**

**Also, similarly, stresses.**

**=> Optimization of geometry (lattice parameters and atomic positions)**

# Density Functional Perturbation Theory

Many physical properties = derivatives of total energy  
(or suitable thermodynamic potential) with respect to perturbations.

Consider :

- atomic displacements (phonons)
- dilatation/contraction of primitive cell
- homogeneous external field (electric field, magnetic field ...)

Derivatives of total energy (electronic part + nuclei-nuclei interaction) :

1<sup>st</sup> order derivatives : forces, stresses, dipole moment ...

2<sup>nd</sup> order derivatives : dynamical matrix, elastic constants, dielectric susceptibility  
atomic polar tensors or Born effective charge tensors  
piezoelectricity, internal strains ...

3<sup>rd</sup> order derivatives : non-linear dielectric susceptibility, Raman susceptibilities  
electro-optic effect, phonon - phonon interaction, Grüneisen parameters, ...

Further properties obtained by integration over phononic degrees of freedom :  
entropy, thermal expansion, phonon-limited thermal conductivity ...

# Iterative algorithms, troubleshooting

# Algorithmics : problems to be solved

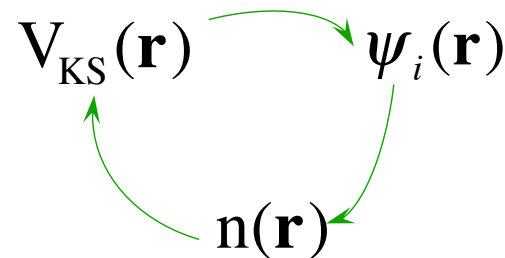
## (1) Kohn - Sham equation

$$\underline{\underline{A}} \underline{x}_i = \lambda_i \underline{x}_i$$

$$\left[ \begin{array}{c} -\frac{1}{2} \nabla^2 + V_{\text{KS}}(\mathbf{r}) \\ \{\mathbf{G}_j\} \quad \{\mathbf{r}_j\} \end{array} \right] \psi_i(\mathbf{r}) = \varepsilon_i \psi_i(\mathbf{r})$$

Size of the system	[2 atoms...	600 atoms...] + vacuum ?
Dimension of the vectors $\underline{x}_i$	300...	100 000... (if planewaves)
# of (occupied) eigenvectors	4...	1200...

## (2) Self-consistency



## (3) Geometry optimization

Find the positions  $\{\mathbf{R}_\kappa\}$  of ions such that the forces  $\{\mathbf{F}_\kappa\}$  vanish  
 [ = Minimization of energy ]

Current practice : iterative approaches

# The 'steepest-descent' algorithm

Forces are gradients of the energy : moving the atoms along gradients is the steepest descent of the energy surface.

=> Iterative algorithm.

Choose a starting geometry, then a parameter  $\lambda$  ,  
and iterately update the geometry, following the forces :

$$\mathbf{R}_{\kappa,\alpha}^{(n+1)} = \mathbf{R}_{\kappa,\alpha}^{(n)} + \lambda \mathbf{F}_{\kappa,\alpha}^{(n)}$$

Equivalent to the simple mixing algorithm  
of SCF (see later)

# Energy+forces around equilb. geometry

Let us denote the equilibrium geometry as  $\mathbf{R}_{\kappa,\alpha}^*$

Analysis of forces close to the equilibrium geometry,  
at which forces vanish, thanks to a Taylor expansion :

$$F_{\kappa,\alpha}(\mathbf{R}_{\kappa',\alpha'}) = \cancel{F_{\kappa,\alpha}(\mathbf{R}_{\kappa',\alpha'}^*)} + \sum_{\kappa',\alpha'} \left. \frac{\partial F_{\kappa,\alpha}}{\partial \mathbf{R}_{\kappa',\alpha'}} \right|_{\{\mathbf{R}^*\}} (\mathbf{R}_{\kappa',\alpha'} - \mathbf{R}_{\kappa',\alpha'}^*) + O(\mathbf{R}_{\kappa',\alpha'} - \mathbf{R}_{\kappa',\alpha'}^*)^2$$

Moreover,  $F_{\kappa,\alpha} = -\frac{\partial E^{\text{BO}}}{\partial \mathbf{R}_{\kappa,\alpha}}$   $\frac{\partial F_{\kappa',\alpha'}}{\partial \mathbf{R}_{\kappa,\alpha}} = -\frac{\partial^2 E^{\text{BO}}}{\partial \mathbf{R}_{\kappa,\alpha} \partial \mathbf{R}_{\kappa',\alpha'}}$

Vector and matrix notation

$$\begin{array}{lll} \mathbf{R}_{\kappa,\alpha}^* \rightarrow \underline{\mathbf{R}}^* & F_{\kappa,\alpha} \rightarrow \underline{\mathbf{F}} & \left. \frac{\partial^2 E^{\text{BO}}}{\partial \mathbf{R}_{\kappa,\alpha} \partial \mathbf{R}_{\kappa',\alpha'}} \right|_{\{\mathbf{R}_{\kappa,\alpha}^*\}} \rightarrow \underline{\underline{\mathbf{H}}} \\ \mathbf{R}_{\kappa,\alpha} \rightarrow \underline{\mathbf{R}} & & \text{(the Hessian)} \end{array}$$

# Steepest-descent : analysis (I)

$$\mathbf{R}_{K,\alpha}^{(n+1)} = \mathbf{R}_{K,\alpha}^{(n)} + \lambda \mathbf{F}_{K,\alpha}^{(n)}$$

Analysis of this algorithm, in the linear regime :

$$\underline{\mathbf{F}}(\underline{\mathbf{R}}) = \cancel{\underline{\mathbf{F}}(\underline{\mathbf{R}}^*)} - \underline{\mathbf{H}}(\underline{\mathbf{R}} - \underline{\mathbf{R}}^*) + \cancel{O(\underline{\mathbf{R}} - \underline{\mathbf{R}}^*)^2}$$

$$\underline{\mathbf{R}}^{(n+1)} = \underline{\mathbf{R}}^{(n)} + \lambda \underline{\mathbf{F}}^{(n)} \longrightarrow (\underline{\mathbf{R}}^{(n+1)} - \underline{\mathbf{R}}^*) = (\underline{\mathbf{R}}^{(n)} - \underline{\mathbf{R}}^*) - \lambda \underline{\mathbf{H}}(\underline{\mathbf{R}}^{(n)} - \underline{\mathbf{R}}^*)$$

$$(\underline{\mathbf{R}}^{(n+1)} - \underline{\mathbf{R}}^*) = (\underline{\mathbf{1}} - \lambda \underline{\mathbf{H}})(\underline{\mathbf{R}}^{(n)} - \underline{\mathbf{R}}^*)$$

For convergence of the iterative procedure, the "distance" between trial geometry and equilibrium geometry must decrease.

- 1) Can we predict conditions for convergence ?
- 2) Can we make convergence faster ?

Need to understand the action of the matrix (or operator)

$$\underline{\mathbf{1}} - \lambda \underline{\mathbf{H}}$$



# Steepest-descent : analysis (II)

What are the eigenvectors and eigenvalues of  $\underline{\underline{H}}$  ?

$\underline{\underline{H}}$  symmetric,  
positive definite matrix

$$\left( = \frac{\partial^2 E^{BO}}{\partial R_{\kappa, \alpha} \partial R_{\kappa', \alpha'}} \bigg|_{\{R_{\kappa, \alpha}^*\}} \right)$$

$\underline{\underline{H}} \underline{f}_i = h_i \underline{f}_i$  where  $\{\underline{f}_i\}$  form a complete, orthonormal, basis set

Discrepancy decomposed as  $(\underline{R}^{(n)} - \underline{R}^*) = \sum_i c_i^{(n)} \underline{f}_i$

and  $(\underline{R}^{(n+1)} - \underline{R}^*) = (\underline{1} - \lambda \underline{\underline{H}}) \sum_i c_i^{(n)} \underline{f}_i = \sum_i c_i^{(n)} (1 - \lambda h_i) \underline{f}_i$

The coefficient of  $\underline{f}_i$  is multiplied by  $1 - \lambda h_i$

Iteratively :

$$(\underline{R}^{(n)} - \underline{R}^*) = \sum_i c_i^{(0)} (1 - \lambda h_i)^{(n)} \underline{f}_i$$

# Steepest-descent : analysis (III)

$$\left(\underline{\mathbf{R}}^{(n)} - \underline{\mathbf{R}}^*\right) = \sum_i c_i^{(0)} (1 - \lambda h_i)^{(n)} \underline{\mathbf{f}}_i$$

The size of the discrepancy decreases if  $|1 - \lambda h_i| < 1$

Is it possible to have  $|1 - \lambda h_i| < 1$  , for all eigenvalues ?

H positive definite  $\Rightarrow$  all  $h_i$  are positive

Yes ! If  $\lambda$  positive, sufficiently small ...

# Steepest-descent : analysis (IV)

$$\left(\underline{\mathbf{R}}^{(n)} - \underline{\mathbf{R}}^*\right) = \sum_i c_i^{(0)} (1 - \lambda h_i)^{(n)} \underline{\mathbf{f}}_i$$

How to determine the optimal value of  $\lambda$  ?

The maximum of all  $|1 - \lambda h_i|$  should be as small as possible.

At the optimal value of  $\lambda$  , what will be the convergence rate ?

( = by which factor is reduced the worst component of  $\left(\underline{\mathbf{R}}^{(n)} - \underline{\mathbf{R}}^*\right)$  ? )

As an exercise : suppose  $\left. \begin{array}{l} h_1 = 0.2 \\ h_2 = 1.0 \\ h_3 = 5.0 \end{array} \right\} \Rightarrow \text{what is the best value of } \lambda \text{ ?}$   
+ what is the convergence rate ?

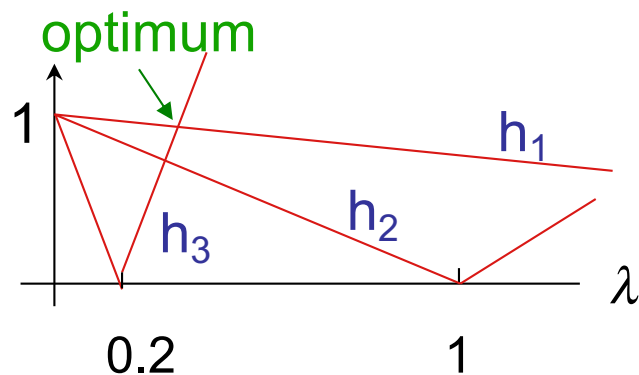
Hint : draw the three functions  $|1 - \lambda h_i|$  as a function of  $\lambda$  . Then, find the location of  $\lambda$  where the largest of the three curves is the smallest.  
Find the coordinates of this point.

# Steepest-descent : analysis (V)

Minimise the maximum of  $|1 - \lambda h_i|$

$$\begin{array}{lcl} h_1 = 0.2 & |1 - \lambda \cdot 0.2| \\ h_2 = 1.0 & |1 - \lambda \cdot 1| \\ h_3 = 5.0 & |1 - \lambda \cdot 5| \end{array}$$

$$\begin{array}{lcl} \text{optimum} \Rightarrow \lambda = 5 \\ \text{optimum} \Rightarrow \lambda = 1 \\ \text{optimum} \Rightarrow \lambda = 0.2 \end{array} \quad \left. \vphantom{\begin{array}{l} \text{optimum} \Rightarrow \lambda = 5 \\ \text{optimum} \Rightarrow \lambda = 1 \\ \text{optimum} \Rightarrow \lambda = 0.2 \end{array}} \right\} ?$$



$$\mu = |1 - \underbrace{\lambda \cdot 0.2}_{\text{positive}}| = |1 - \underbrace{\lambda \cdot 5}_{\text{negative}}|$$

$$1 - \lambda \cdot 0.2 = -(1 - \lambda \cdot 5)$$

$$2 - \lambda (0.2 + 5) = 0 \Rightarrow \lambda = 2/5.2$$

$$\mu = 1 - 2 \cdot (0.2 / 5.2)$$

Only ~ 8% decrease of the error, per iteration ! Hundreds of iterations will be needed to reach a reduction of the error by 1000 or more.

Note : the second eigenvalue does not play any role.

The convergence is limited by the extremal eigenvalues : if the parameter is too large, the smallest eigenvalue will cause divergence, but for that small parameter, the largest eigenvalue lead to slow decrease of the error...

# The condition number

In general,  $\lambda_{\text{opt}} = 2 / (h_{\text{min}} + h_{\text{max}})$   
 $\mu_{\text{opt}} = 2 / [1 + (h_{\text{max}}/h_{\text{min}})] - 1 = [(h_{\text{max}}/h_{\text{min}}) - 1] / [(h_{\text{max}}/h_{\text{min}}) + 1]$

Perfect if  $h_{\text{max}} = h_{\text{min}}$ . Bad if  $h_{\text{max}} \gg h_{\text{min}}$ .

$h_{\text{max}}/h_{\text{min}}$  called the "condition" number. A problem is "ill-conditioned" if the condition number is large. It does **not** depend on the intermediate eigenvalues.

Suppose we start from a configuration with forces on the order of 1 Ha/Bohr, and we want to reach the target  $1e-4$  Ha/Bohr. The mixing parameter is optimal.

How many iterations are needed ?

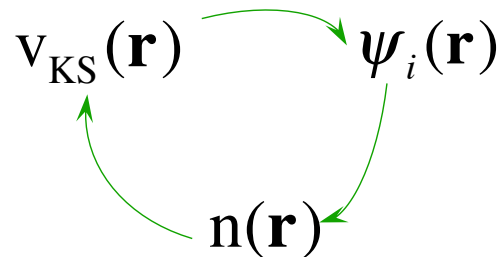
For a generic decrease factor  $\Delta$ , with "n" the number of iterations.

$$\|\underline{F}^{(n)}\| \approx \left( \frac{h_{\text{max}}/h_{\text{min}} - 1}{h_{\text{max}}/h_{\text{min}} + 1} \right)^n \|\underline{F}^{(0)}\| \quad \Delta \approx \left( \frac{h_{\text{max}}/h_{\text{min}} - 1}{h_{\text{max}}/h_{\text{min}} + 1} \right)^n$$

$$n \approx \left[ \ln \left( \frac{h_{\text{max}}/h_{\text{min}} + 1}{h_{\text{max}}/h_{\text{min}} - 1} \right) \right]^{-1} \ln \Delta \approx 0.5 (h_{\text{max}}/h_{\text{min}}) \ln \frac{1}{\Delta}$$

(The latter approximate equality supposes a large condition number)

# Analysis of self-consistency



Natural iterative methodology (KS : in => out) :

$$\mathbf{v}_{\text{in}}(\mathbf{r}) \rightarrow \psi_i(\mathbf{r}) \rightarrow \mathbf{n}(\mathbf{r}) \rightarrow \mathbf{v}_{\text{out}}(\mathbf{r})$$

Which quantity plays the role of a force, that should vanish at the solution ?

The difference

$$\mathbf{v}_{\text{out}}(\mathbf{r}) - \mathbf{v}_{\text{in}}(\mathbf{r}) \quad (\text{generic name : a "residual"})$$

Simple mixing algorithm  
( $\approx$  steepest - descent )

$$\underline{\mathbf{v}}_{\text{in}}^{(n+1)} = \underline{\mathbf{v}}_{\text{in}}^{(n)} + \lambda \left( \underline{\mathbf{v}}_{\text{out}}^{(n)} - \underline{\mathbf{v}}_{\text{in}}^{(n)} \right)$$

Analysis ...  $\underline{\mathbf{v}}_{\text{out}} \left[ \underline{\mathbf{v}}_{\text{in}} \right] = \underline{\mathbf{v}}_{\text{out}} \left[ \underline{\mathbf{v}}^* \right] + \frac{\delta \underline{\mathbf{v}}_{\text{out}}}{\delta \underline{\mathbf{v}}_{\text{in}}} \left( \underline{\mathbf{v}}_{\text{in}} - \underline{\mathbf{v}}^* \right)$

$\delta \underline{\mathbf{v}}_{\text{in}}$   $\rightarrow$   $\underline{\underline{\mathbf{H}}}$

Like the steepest-descent algorithm, this leads to the requirement to minimize  $|1 - \lambda h_i|$  where  $h_i$  are eigenvalues of

$$\frac{\delta \underline{\mathbf{v}}_{\text{out}}}{\delta \underline{\mathbf{v}}_{\text{in}}}$$

(actually, the dielectric matrix)

AMM lecture 21 April Part A

# Modify the condition number (II)

$$\underline{\underline{\mathbf{R}}}^{(n+1)} = \underline{\underline{\mathbf{R}}}^{(n)} + \lambda \left( \underline{\underline{\mathbf{H}}}^{-1} \right)_{\text{approx}} \underline{\underline{\mathbf{F}}}^{(n)}$$

$$\underline{\underline{\mathbf{F}}}(\underline{\underline{\mathbf{R}}}) = -\underline{\underline{\mathbf{H}}}(\underline{\underline{\mathbf{R}}} - \underline{\underline{\mathbf{R}}}^*) \implies (\underline{\underline{\mathbf{R}}}^{(n+1)} - \underline{\underline{\mathbf{R}}}^*) = \left( 1 - \lambda \left( \underline{\underline{\mathbf{H}}}^{-1} \right)_{\text{approx}} \underline{\underline{\mathbf{H}}} \right) (\underline{\underline{\mathbf{R}}}^{(n)} - \underline{\underline{\mathbf{R}}}^*)$$

- Notes :
- 1) If approximate inverse Hessian perfect, optimal geometry is reached in one step, with  $\lambda = 1$ .  
Steepest-descent NOT the best direction.
  - 2) Non-linear effects not taken into account. For geometry optimization, might be quite large. Even with perfect hessian, need 5-6 steps to optimize a water molecule.
  - 3) Approximating inverse hessian by a multiple of the unit matrix is equivalent to changing the  $\lambda$  value.
  - 4) Eigenvalues and eigenvectors of  $\left( \underline{\underline{\mathbf{H}}}^{-1} \right)_{\text{approx}} \underline{\underline{\mathbf{H}}}$  govern the convergence : the condition number can be changed.  
 $\left( \underline{\underline{\mathbf{H}}}^{-1} \right)_{\text{approx}}$  often called a "**pre-conditioner**".
  - 5) Generalisation to other optimization problems is trivial.  
(The Hessian is referred to as the Jacobian if it is not symmetric.)

# Modify the condition number (III)

Approximate Hessian can be generated on a case-by-case basis.

Selfconsistent determination of the Kohn-Sham potential :

Jacobian = dielectric matrix.

Lowest eigenvalue **close to 1**.

Largest eigenvalue :

= **1.5 ... 2.5** for small close-shell molecules, and small unit cell solids

(Simple mixing will sometimes converge with parameter set to 1 !)

= the **macroscopic dielectric constant** (e.g. 12 for silicon),

for larger close-shell molecules and large unit cell insulators,

= **diverge** for large-unit cell metals, or open-shell molecules !

**Model** dielectric matrices known for rather homogeneous systems.

Knowledge of approx. macroscopic dielectric constant

=> efficient preconditioner

Work in progress for inhomogeneous systems

(e.g. metals/vacuum systems).



# Advanced algorithms : using the history

Instead of using only previously computed forces,  
take into account past forces for past positions

Large class of methods :

- Broyden (quasi-Newton-type),
- Davidson,
- conjugate gradients,
- Lanczos ...

+Approximate Hessian can be combined with usage of history

For **geometry optimization** : current practice

Broyden **ionmov 22**

First without cell relaxation **optcell 0**

then optimizing the cell **optcell 2** (or **1**, or **3...9**)

# Iterative algorithms always working ?

No !

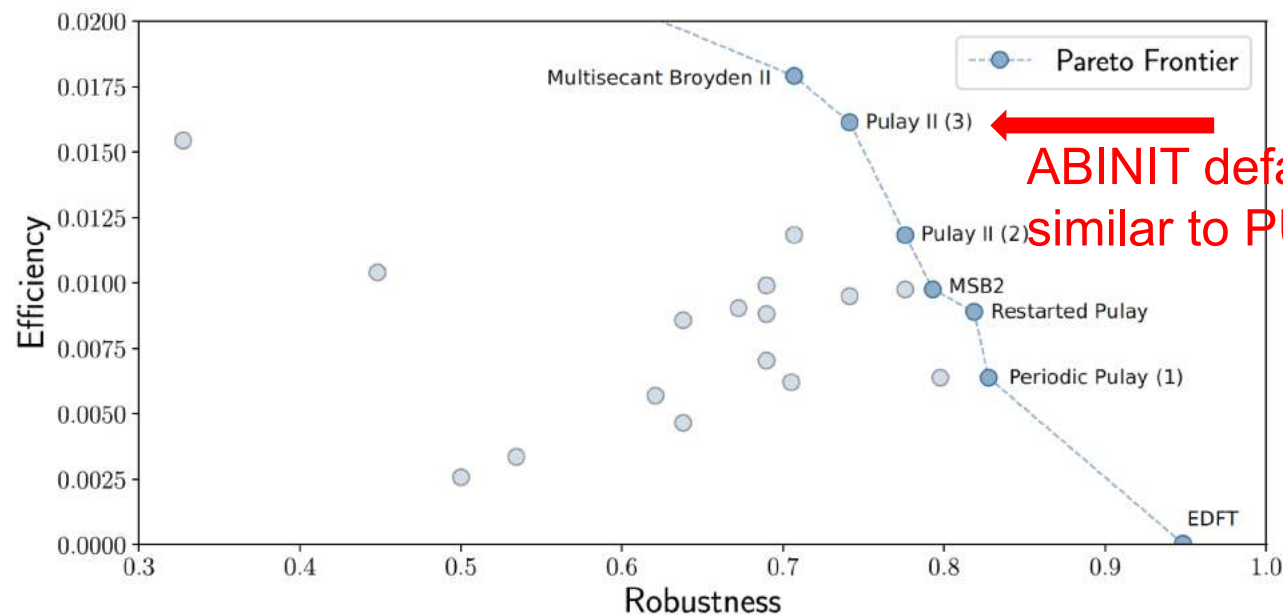
Hypothesis of harmonic energy can be problematic

Many hard cases where iterations go on, without reaching SCF convergence or vanishing forces / stresses.

Trade-off between speed and robustness.

Example : test of 24 iterative algorithms/settings for a set of 56 hard systems

N.D. Woods, M.C. Payne and P.J. Hasnip, *J. Phys.: Condens. Matt.* **31**, 453001 (2019)



ABINIT default algo+settings  
similar to PULAY II (3)

# Iterative algorithms always working ?

SCF default **iscf** 7 or 17

try changing **iscf**, **diemac**, **dielng**

is the solution of KS equation accurate ? **nline**, **nnscl**, **tolrde**

Geometry optimization default **ionmov** 22

try improving underlying SCF accuracy using **tolrff**, **nstep**

try changing **ionmov**

if optimizing primitive lattice vectors, beware about **ecutsm**,

# Other troubleshooting ?

Read ABINIT error messages !

Possibly also warnings and comments.

In log file, or in error file, or in the `__ABI_MPIABORTFILE__` file if parallel

Problems with the initial geometry (cell, atom type, nuclei positions)

**Many** mistakes done by beginners !

Check the units (reminder : ABINIT uses atomic units by default)

Check whether natom is coherent with xred (or xcart)

Check that you have not switched atom types

Relax first the atomic positions **at fixed primitive vectors** before optimizing the cell. Read optcell documentation

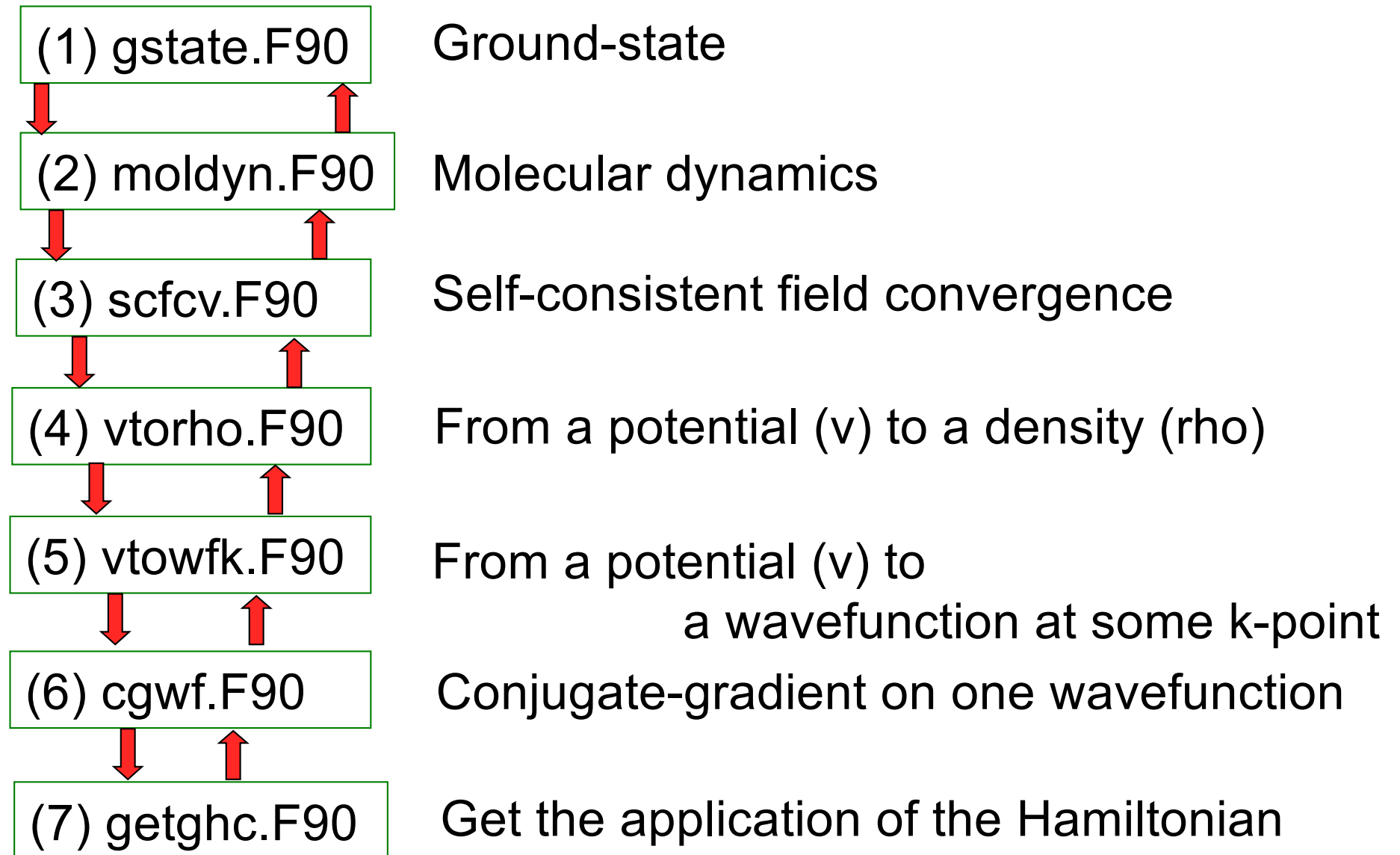
ABINIT can read VASP POSCAR external files containing unit cell parameters and atomic positions. See the input variable structure.

Possibly try to visualize your input geometry using Abipy, VESTA, XCrysDEN, see [docs.abinit.org/tutorial/analysis\\_tools](https://docs.abinit.org/tutorial/analysis_tools) to produce .xsf files for XCrysDEN

More on troubleshooting at

[wiki.abinit.org/doku.php?id=howto:troubleshooting](https://wiki.abinit.org/doku.php?id=howto:troubleshooting)

# ABINIT : levels in main processing unit



# An example run : tbase1\_1

- First run in tutorial base1 (<https://docs.abinit.org/tutorial/base1>)
- Very quick (less than 1 sec)
- Look at tbase1\_1.files, tbase1\_1.in, 01h.pspgth
- `abinit < tbase1_1.files > log`
- Look at tbase1\_1.out (such \*.out file will usually be much longer)
- Look at log (such log file is mainly for finding errors or debugging )
- Issue: (-A3 might not be enough for more atoms)

```
grep -i -B1 -A3 ETOT tbase1_1.out
```

```
grep -i -A3 forces      tbase1_1.out
```

```
grep -i -A3 stress     tbase1_1.out
```

```
grep -i -A3 energy     tbase1_1.out
```

```
grep -i -A3 coordinates tbase1_1.out
```

```
grep -i -A3 density    tbase1_1.out
```

# Additional slides

# Exact result for exchange-correlation energy

(without demonstration)

The exchange-correlation energy, functional of the density is the integral over the whole space of the density times the local exchange-correlation energy per particle

$$E_{xc}[n] = \int n(\mathbf{r}_1) \varepsilon_{xc}(\mathbf{r}_1; n) d\mathbf{r}_1$$

while the local exchange-correlation energy per particle is the electrostatic interaction energy of a particle with its DFT exchange-correlation hole.

$$\varepsilon_{xc}(\mathbf{r}_1; n) = \int \frac{1}{2} \frac{\overline{n^{xc}}(\mathbf{r}_2 | \mathbf{r}_1; n)}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_2$$

$$\text{Sum rule : } \int \overline{n^{xc}}(\mathbf{r}_2 | \mathbf{r}_1; n) d\mathbf{r}_2 = -1$$



# Local density approximation (I)

Hypothesis :

- the local XC energy per particle only depend on the local density
- and is equal to the local XC energy per particle of an homogeneous electron gas of same density (in a neutralizing background - « jellium » )

$$\epsilon_{\text{xc}}^{\text{LDA}}(\mathbf{r}_1; n) = \epsilon_{\text{xc}}^{\text{hom}}(n(\mathbf{r}_1))$$

Gives excellent numerical results ! Why ?

- 1) Sum rule is fulfilled
- 2) Characteristic screening length indeed depend on local density

# Local density approximation (II)

Actual function : exchange part + correlation part

$$\varepsilon_x^{\text{hom}}(n) = Cn^{1/3} \quad \text{with} \quad C = -\frac{3}{4\pi}(3\pi^2)^{1/3}$$

for the correlation part, one resorts to accurate numerical simulations beyond DFT (e.g. Quantum Monte Carlo)

Corresponding exchange-correlation potential  $V_{\text{xc}}(\mathbf{r}) = \frac{\delta E_{\text{xc}}[n]}{\delta n(\mathbf{r})}$

$$V_{\text{xc}}^{\text{approx}}(\mathbf{r}) = \mu_{\text{xc}}(n(\mathbf{r})) \quad \mu_{\text{xc}}(n) = \frac{d(n\varepsilon_{\text{xc}}^{\text{approx}}(n))}{dn}$$

$$\mu_x(n) = C \frac{4}{3} n^{1/3} = \frac{4}{3} \varepsilon_x^{\text{hom}}(n)$$

# Hund's rule

Example: p levels

	$\uparrow \text{ --- } \text{ --- }$	$\uparrow \uparrow \text{ --- }$	$\uparrow \uparrow \uparrow$	$\uparrow\downarrow \uparrow \uparrow$	$\uparrow\downarrow \uparrow\downarrow \uparrow$	$\uparrow\downarrow \uparrow\downarrow \uparrow\downarrow$
$ S _{\text{total}}$	1/2	1	3/2	1	1/2	0

Suppose frozen spherically symmetric effective potential  
created by nuclei and core electrons

$\uparrow \uparrow \text{ --- }$  favored over  $\uparrow \downarrow \text{ --- }$

Indeed, density identical => Hartree energy identical  
However, exchange energy more negative

$$E_H[n] = \frac{1}{2} \int \frac{n(\mathbf{r}_1)n(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2 \quad E_x = -\frac{1}{2} \sum_{\sigma} \sum_{i,j}^{occ} \int d\mathbf{r} d\mathbf{r}' \frac{\psi_{j\sigma}^*(\mathbf{r}') \psi_{i\sigma}^*(\mathbf{r}') \psi_{j\sigma}(\mathbf{r}) \psi_{i\sigma}(\mathbf{r})}{|\mathbf{r} - \mathbf{r}'|}$$

$\uparrow \uparrow \text{ --- }$  favored over  $\uparrow\downarrow \text{ --- } \text{ --- }$

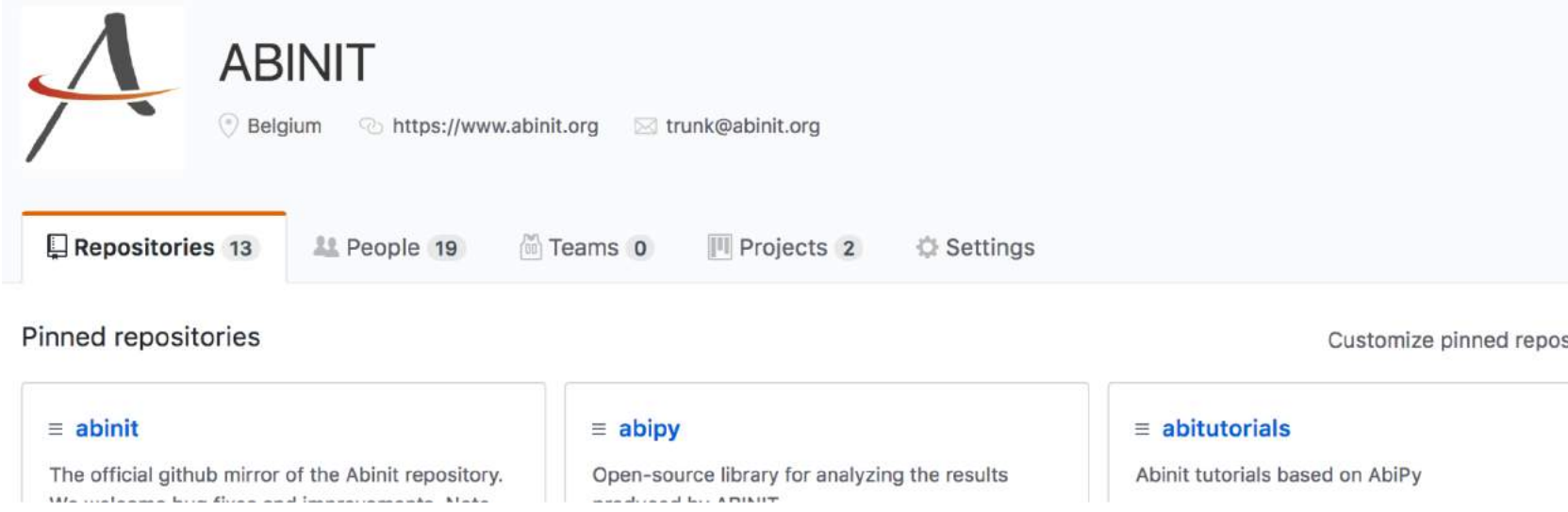
Indeed, density in the same orbital (same spatial location)  
increase Hartree energy (also  $E_x$  energy less negative !)

# ABINIT + python : Abipy, Abitutorials ...

ABINIT organization on GitHub <https://github.com/abinit>

**Abipy** : python library for launching ABINIT jobs,  
and analysing/plotting the results <http://pythonhosted.org/abipy>  
=> e.g. connecting ABINIT with tools for high-throughput  
calculations developed in the **Materials Project** context  
(like Pymatgen, Fireworks).

**Abitutorials** : tutorial based on Jupyter notebooks ABINIT+python



The screenshot shows the GitHub profile of the ABINIT organization. At the top, there is a logo consisting of a stylized 'A' with a red and orange swoosh. To the right of the logo, the name 'ABINIT' is displayed, followed by the location 'Belgium', the website 'https://www.abinit.org', and the email 'trunk@abinit.org'. Below this, a navigation bar shows 'Repositories 13', 'People 19', 'Teams 0', 'Projects 2', and 'Settings'. The 'Pinned repositories' section is visible, showing three pinned repositories: 'abinit' (The official github mirror of the Abinit repository), 'abipy' (Open-source library for analyzing the results), and 'abitutorials' (Abinit tutorials based on AbiPy).

# Advanced Materials Modeling

## Lecture of 21 april

### Part B

- + Many-body physics
- + The total energy : Quantum monte Carlo
- + Experiments : total energy, band gap, spectroscopies ?
- + Charged excitations and Green's function  $G$
- + GW approximation (incl. self-consistency and miscellaneous tricks)

# A basic reference on first-principles simulations of Interacting electrons

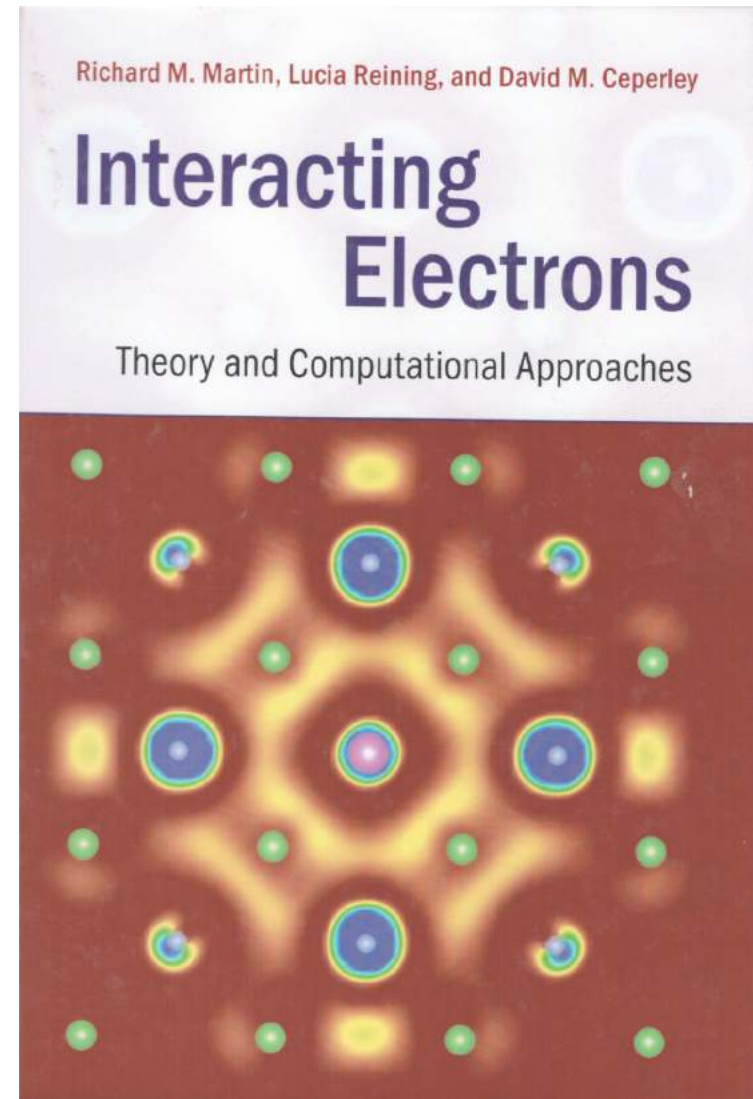
Richard M. Martin, L. Reining, and D.M.  
Ceperley

Cambridge University Press, 2016

Interacting electrons:

Theory and computational approaches

(ISBN: 978-0-521)87150-1)



# Many-body physics

# Systems with n electrons and N nuclei

- Born-Oppenheimer approximation.
- Nuclei are treated as distinct particles, without spin. Internal structure is neglected. Only charge, mass and (classical) position taken into account.

Wavefunction :

$$\Psi_e(\mathbf{r}_1, \sigma_1, \mathbf{r}_2, \sigma_2, \dots, \mathbf{r}_n, \sigma_n; \mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_N) \quad + \text{antisymmetry for electron exchange}$$

Electronic Hamiltonian, in atomic units (without spin-orbit) :

$$\hat{T}_e = \sum_{i=1}^n -\frac{\nabla_{\mathbf{r}_i}^2}{2} \quad \hat{V}_{eN} = \sum_A^N \sum_i^n \frac{-Z_A}{|\mathbf{r}_i - \mathbf{R}_A|} \quad \hat{V}_{ee} = \sum_{\substack{(i,j) \\ i < j}}^{(n,n)} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}$$
$$\hat{H}_e = \hat{T}_e + \hat{V}_{eN} + \hat{V}_{ee}$$

$$\hat{H}_e |\Psi_e\rangle = E_e |\Psi_e\rangle$$



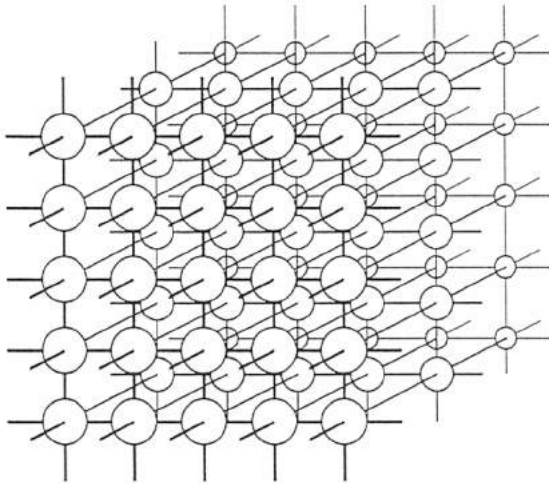
# Many-body wavefunction representation

Classical objects : fields  
or trajectories

$$p(\mathbf{r},t), V(\mathbf{r},t), T(\mathbf{r},t), \mathbf{E}(\mathbf{r},t), \dots$$
$$\mathbf{R}_1(t), \mathbf{R}_2(t), \mathbf{R}_3(t), \dots, \mathbf{R}_N(t)$$

Quantum objects : wavefunctions  
for interacting particles

$$\Psi(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \dots, \mathbf{r}_N, t)$$



Classical position and speed of 8 objects  
 $2 \times 3 \times 8 = 48$  real numbers.

Oxygen atom : 8 electrons.

Quantum description, on a cubic grid of  $10 \times 10 \times 10$  points.

24-dimensional object  $\Rightarrow 10^{24}$  complex numbers

HF, DFT : set of wavefunctions for non-interacting particles

For the oxygen atom, back to  $8 \times 10 \times 10 \times 10$  real numbers,  
but with an approximate treatment ...

# Simplified notations

(1) Arguments of wavefunction :

$$\Psi_e(\mathbf{r}_1, \sigma_1; \mathbf{r}_2, \sigma_2; \dots; \mathbf{r}_n, \sigma_n) \rightarrow \Psi_e(1, 2, \dots, n)$$

“1” becomes a symbol to denote  $\mathbf{r}_1, \sigma_1$

(2) Define also  $r_{ij} \triangleq |\mathbf{r}_i - \mathbf{r}_j|$   $\hat{V}_{ee} = \sum_{\substack{(i,j) \\ i < j}}^{(n,n)} \frac{1}{r_{ij}}$

and “one-particle” part of Hamiltonian

$$\hat{V}_{eN} = \sum_A^N \sum_i^n \frac{-Z_A}{r_{iA}} = \sum_i^n \hat{V}_N(\mathbf{r}_i) \quad \hat{h}^N(i) \triangleq -\frac{\nabla_{r_i}^2}{2} + \hat{V}_N(\mathbf{r}_i)$$

Hamiltonian becomes :

$$\left( \sum_i^n \hat{h}^N(i) + \sum_{\substack{(i,j) \\ i < j}}^{(n,n)} \frac{1}{r_{ij}} \right) \Psi_e(1, 2, \dots, n) = E_e \Psi_e(1, 2, \dots, n)$$

# Separation of variables

$$\left( \sum_i^n \hat{h}^N(i) + \sum_{\substack{(i,j) \\ i < j}}^{(n,n)} \frac{1}{r_{ij}} \right) \Psi_e(1,2,\dots,n) = E_e \Psi_e(1,2,\dots,n)$$

+ antisymmetry for electron exchange

Interelectronic coupling does not  
allow to separate electronic coordinates

Mean-field approximation (like DFT)

$$\sum_{\substack{(i,j) \\ i < j}}^{(n,n)} \frac{1}{r_{ij}} \text{ replaced by } \sum_i \hat{V}_{Hxc}(i) \quad \hat{h}_{MF}^N(i) = \hat{h}^N(i) + \hat{V}_{Hxc}(i)$$

$$\left( \sum_i^n \hat{h}_{MF}^N(i) \right) \Psi_{MF,e}(1,2,\dots,n) = E_{MF} \Psi_{MF,e}(1,2,\dots,n)$$

+ antisymmetry for electron exchange

gives solutions as

Slater determinants of 1-particle wavefunctions

mean-field energy=sum of 1-particle eigenenergies,

must be corrected to give electronic energy (avoid double-counting)

# Monte Carlo approaches

$$\left( \sum_i^n \hat{h}^N(i) + \sum_{\substack{(i,j) \\ i < j}}^{(n,n)} \frac{1}{r_{ij}} \right) \Psi_e(1,2,\dots,n) = E_e \Psi_e(1,2,\dots,n)$$

Variational Monte Carlo

Diffusion Monte Carlo

Diagrammatic Monte Carlo

Green's function Monte Carlo

...

“Monte Carlo” refers to stochastic evaluation of high-dimensional integrals using random sampling, and generation of such samplings using random walks

Usually : target evaluation of the **ground-state energy**

# Monte Carlo approaches (brief overview)

# Variational Monte Carlo (VMC)

$$\hat{H}_e \Psi_e(1,2,\dots,n) = \left( \sum_i \hat{h}^N(i) + \sum_{\substack{(i,j) \\ i < j}}^{(n,n)} \frac{1}{r_{ij}} \right) \Psi_e(1,2,\dots,n) = E_e \Psi_e(1,2,\dots,n)$$

Use the variational principle to approximate ground-state energy

$$E_{e,GS} = \min_{\Psi_e \text{ normalized}} \langle \Psi_e | \hat{H}_e | \Psi_e \rangle < E_{e,trial} = \langle \Psi_{e,trial} | \hat{H}_e | \Psi_{e,trial} \rangle$$

Simplest form of VMC : use trial wavefunctions of the type

$$\Psi_{e,trial}(1,2,\dots,n) = \left( \prod_{\substack{(i,j) \\ i < j}}^{(n,n)} f(r_{ij}) \right) \Psi_{SD}(1,2,\dots,n)$$

**Jastrow-Slater**

Jastrow factor, symmetric upon electron exchange, builds-in correlation between electrons  
 $f$  tends to 1 at large separation, and to zero at small separation

Slater determinant, antisymmetric upon electron exchange

# Challenge : multi-dimensional integrals

$$\left\langle \Psi_{e,trial} \left| \hat{H}_e \right| \Psi_{e,trial} \right\rangle \quad \text{with} \quad \hat{H}_e = \hat{T}_e + \hat{V}_{eN} + \hat{V}_{ee}$$

$$\hat{T}_e = \sum_{i=1}^n -\frac{\nabla_{r_i}^2}{2} \quad \hat{V}_{eN} = \sum_A^N \sum_i^n \frac{-Z_A}{|\mathbf{r}_i - \mathbf{R}_A|} \quad \hat{V}_{ee} = \sum_{\substack{(i,j) \\ i < j}}^{(n,n)} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}$$

$$\text{and} \quad \Psi_e(1,2,\dots,n) = \left( \prod_{\substack{(i,j) \\ i < j}}^{(n,n)} f(r_{ij}) \right) \Psi_{SD}(1,2,\dots,n)$$

$$\left\langle \Psi_{e,trial} \left| \hat{H}_e \right| \Psi_{e,trial} \right\rangle = \int \dots \int \Psi_{e,trial}^*(1,\dots,N) \hat{H}_e \Psi_{e,trial}(1,\dots,N) d1 \dots dN$$

No simple expression for the expectation value,  
unlike with Slater determinants

Evaluation of Jastrow-Slater wavefunction energy expectation,  
then optimization of orbitals of Slater determinant and Jastrow factor ?

# Stochastic evaluation

$$\left\langle \Psi_{e,trial} \left| \hat{H}_e \right| \Psi_{e,trial} \right\rangle = \int \dots \int \Psi_{e,trial}^*(1, \dots, N) \hat{H}_e \Psi_{e,trial}(1, \dots, N) d1 \dots dN$$

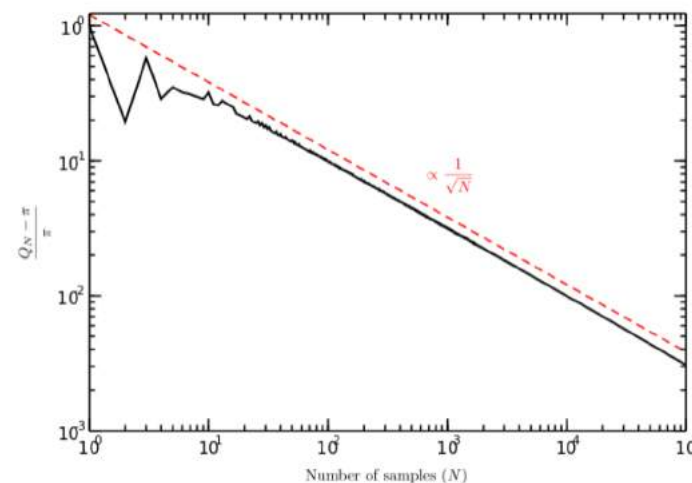
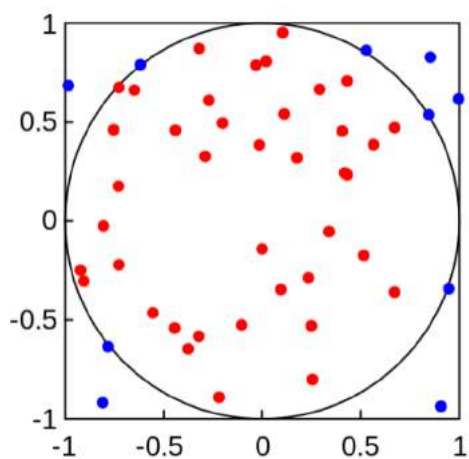
Standard multi (few) dimensional integration of dimension  $d$  :  
sample with homogeneous grid of  $L$  points in each dimension,  
with step  $h$  inversely proportional to  $L$  .

Total number of points  $M = L^d$

Convergence of Simpson rule error  $\varepsilon = (1 / L)^4 = M^{-4/d}$

Random sampling  $\varepsilon = M^{-1/2}$

Better than fixed grid  
technique for  $d > 8$  !





# Importance sampling

Sampling point distribution : generated from the trial wavefunction, to favour the configurations that have higher probability

$$\Pi(1\dots N) = \frac{|\Psi_{e,trial}(1\dots N)|^2}{\int |\Psi_{e,trial}(1\dots N)|^2 d1\dots dN}$$

Metropolis algorithm to generate a random walk in the multi-dimensional space  $(1\dots N)_i$  for  $i=1 \dots M$ ,  
whose probability distribution tends to  $\Pi(1\dots N)$   
by randomly generating new configurations and accepting/rejecting them

Then evaluate  $\langle \Psi_{e,trial} | \hat{H}_e | \Psi_{e,trial} \rangle = \lim_{M \rightarrow \infty} \frac{1}{M} \sum_{i=1}^M E_L((1\dots N)_i)$

where  $E_L((1\dots N)) = \text{Re} \frac{\hat{H}_e \Psi_{e,trial}(1\dots N)}{\Psi_{e,trial}(1\dots N)}$

# Diffusion Monte Carlo

Consider time-dependent Schrödinger equation in imaginary time

$$i\hbar \frac{\partial}{\partial t} |\Psi_e(t)\rangle = \hat{H}_e |\Psi_e(t)\rangle \quad \Rightarrow \quad -\hbar \frac{\partial}{\partial \tau} |\Psi_e(\tau)\rangle = \hat{H}_e |\Psi_e(\tau)\rangle$$

Propagate initial trial wavefunction  $|\Psi_{e,trial}(\tau=0)\rangle$  using such Eq.

Complete set of eigenfunctions of Hamiltonian  $\hat{H}_e |\Psi_i\rangle = E_i |\Psi_i\rangle$

Solve imaginary-time Schrödinger equation  $|\Psi_i(\tau)\rangle = e^{-E_i\tau} |\Psi_i\rangle$

Decomposition of  $|\Psi_{e,trial}(\tau=0)\rangle = \sum_{i=1} c_i |\Psi_i\rangle$

$$|\Psi_{e,trial}(\tau)\rangle = \sum_{i=1} c_i e^{-E_i\tau} |\Psi_i\rangle$$

Hence, in the large-time limit  $|\Psi_{e,trial}(\tau)\rangle \rightarrow |\Psi_{e,i=\text{Ground State}}\rangle$

Diffusion Monte-Carlo = generate a random walk using imaginary-time Schrödinger equation

# Accuracy of Quantum Monte Carlo

Variational Monte Carlo :

quality of Energy estimate depends directly on trial wavefunction

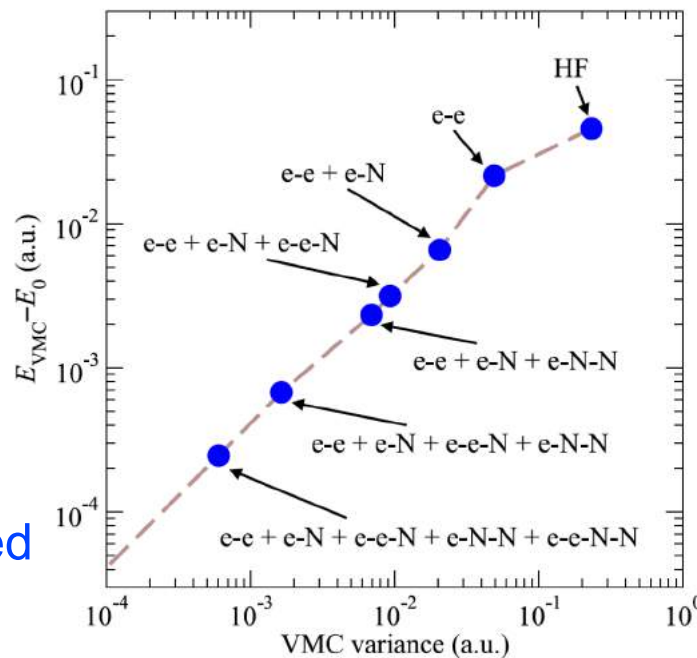
Diffusion Monte Carlo :

quality of Energy is **partly** dependent on trial wavefunction

**Still not exact !**

Indeed, the antisymmetry of the sampling of the random walk is hard to maintain => fixed-node approximation (and fixed-node error)

VMC  $H_2$   
benchmark :  
total  
energy  
difference  
wrt  
numerically  
exact results  
With improved  
Jastrow



## NiO lattice parameter, cohesive energy

J. Phys.: Condens. Matter **30** (2018) 195901

**Table 3.** Values of lattice constant ( $a$ ), bulk modulus ( $B_0$ ), and cohesive energy ( $E_{\text{coh}}$ ) for AFM-II type NiO obtained from a Vinet fit of the equation of state computed using GGA + U and DMC at  $U = U_{\text{opt}}$  and a 16-atom type II AFM NiO supercell.

Method	$a$ (Å)	$B_0$ (GPa)	$E_{\text{coh}}$ (eV/f.u.)
GGA + U	4.234	192	8.54
DMC	4.161(7)	218(14)	9.54(5)
Experiment <sup>a</sup>	4.17	145–206	9.5

# Quantum Monte Carlo codes

## QMCPACK

<https://qmcpack.org>

Jeongnim Kim et al , J. Phys.: Condens. Matter 30, 195901 (2018)

## CASINO

<https://vallico.net/casinoqmc>

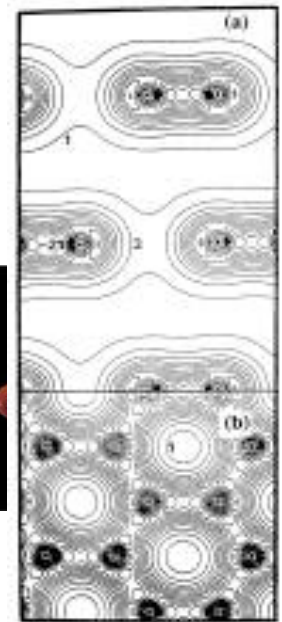
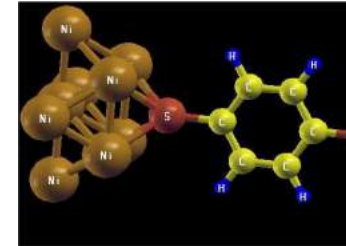
RJ Needs, MD Towler, ND Drummond and P Lopez Rios  
J. Phys.: Condens. Matter 22, 195901 (2010)

# Experiments: total energy, band gap, spectroscopies ?

# Properties from first-principles

Computation of ...

interatomic distances, angles, **total energies**  
electronic charge densities, **electronic energies**



A basis for the computation of ...

chemical reactions

**electronic transport**

vibrational properties

thermal capacity

dielectric behaviour

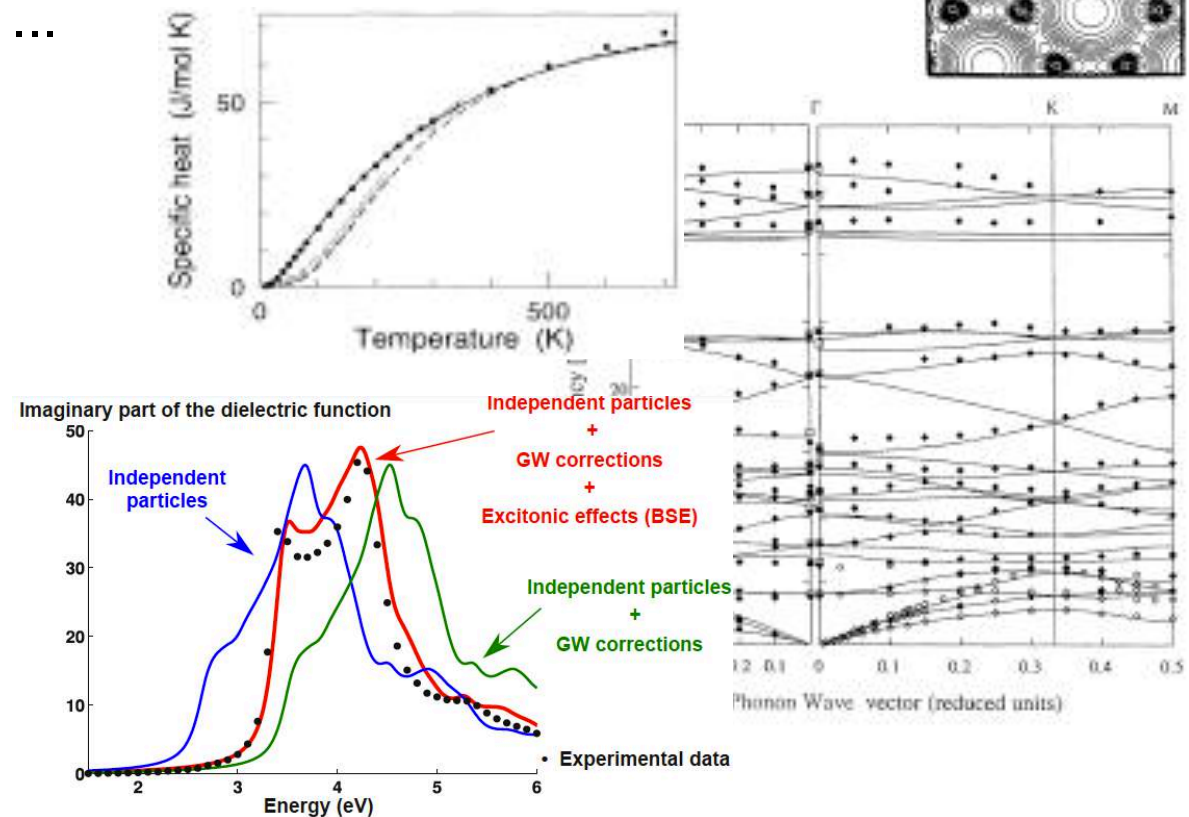
**optical response**

superconductivity

surface properties

**spectroscopic responses**

...



# Density Functional Perturbation Theory

Many physical properties = derivatives of **total energy**  
(or suitable thermodynamic potential) with respect to perturbations.

Consider :

- atomic displacements (phonons)
- dilatation/contraction of primitive cell
- homogeneous external field (electric field, magnetic field ...)

Derivatives of total energy (electronic part + nuclei-nuclei interaction) :

**1<sup>st</sup> order derivatives** : forces, stresses, dipole moment ...

**2<sup>nd</sup> order derivatives** : dynamical matrix, elastic constants, dielectric susceptibility  
atomic polar tensors or Born effective charge tensors  
piezoelectricity, internal strains ...

**3<sup>rd</sup> order derivatives** : non-linear dielectric susceptibility, Raman susceptibilities  
electro-optic effect, phonon - phonon interaction, Grüneisen parameters, ...

Further properties obtained by integration over phononic degrees of freedom :  
entropy, thermal expansion, phonon-limited thermal conductivity ...

# The DFT bandgap problem (I)

- DFT is a ground state theory  
=>no direct interpretation of Kohn-Sham eigenenergies  $\varepsilon_i$  in

$$\left( -\frac{1}{2}\nabla^2 + V_{ext}(\mathbf{r}) + \int \frac{n(\mathbf{r}_1)}{|\mathbf{r}_1 - \mathbf{r}|} d\mathbf{r}_1 + V_{xc}(\mathbf{r}) \right) \psi_i(\mathbf{r}) = \varepsilon_i \psi_i(\mathbf{r})$$

- However  $\{\varepsilon_i\}$  are similar to “quasi-particle” band structure :  
LDA / GGA results for valence bands are rather accurate ...  
but **NOT** for the band gap  $E_g^{KS} = \varepsilon_c - \varepsilon_v$
- The band gap can alternatively be obtained from total energy differences

$$E_g = \{E(N+1) - E(N)\} - \{E(N) - E(N-1)\} \quad [\text{correct expression !}]$$

(where  $E(N)$  is the total energy of the  $N$  - electron system)

Difference between ionisation potential and electronic affinity



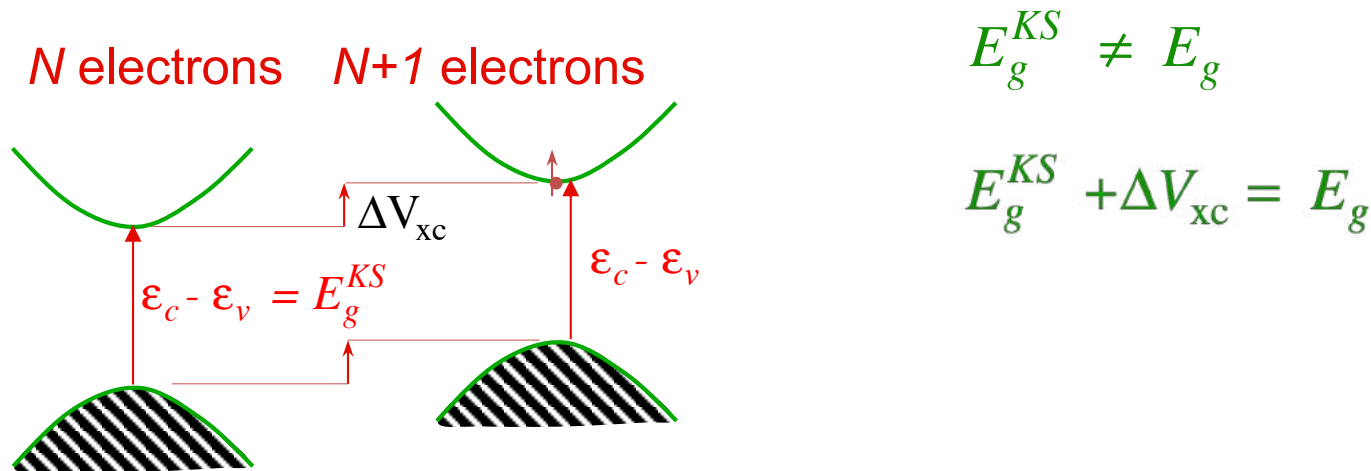
# The DFT bandgap problem (II)

- For LDA & GGA, XC potential is continuous functional of electron number

$$\epsilon_i = \frac{\partial E}{\partial f_i} \quad [\text{Janak's theorem}]$$

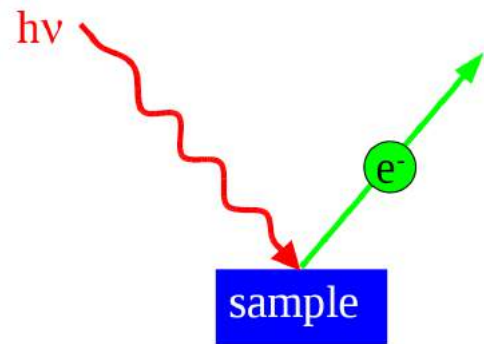
$$\Rightarrow E_g^{KS} = \epsilon_c - \epsilon_v \underset{N \rightarrow \infty}{=} E_g = E(N+1) + E(N-1) - 2E(N)$$

- XC potential : discontinuous with particle number at gap energy

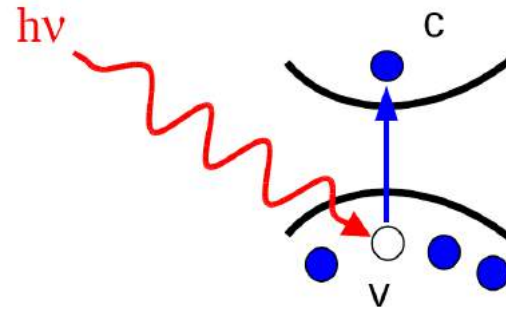


- Band gap : adding / removing an electron = **charged electronic excitations**  
NOT transferring an electron from valence to conduction

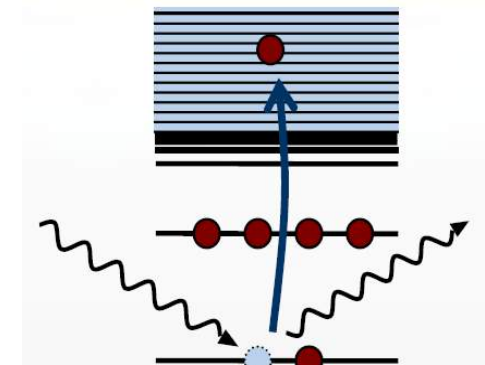
# Spectroscopies : photons & electrons



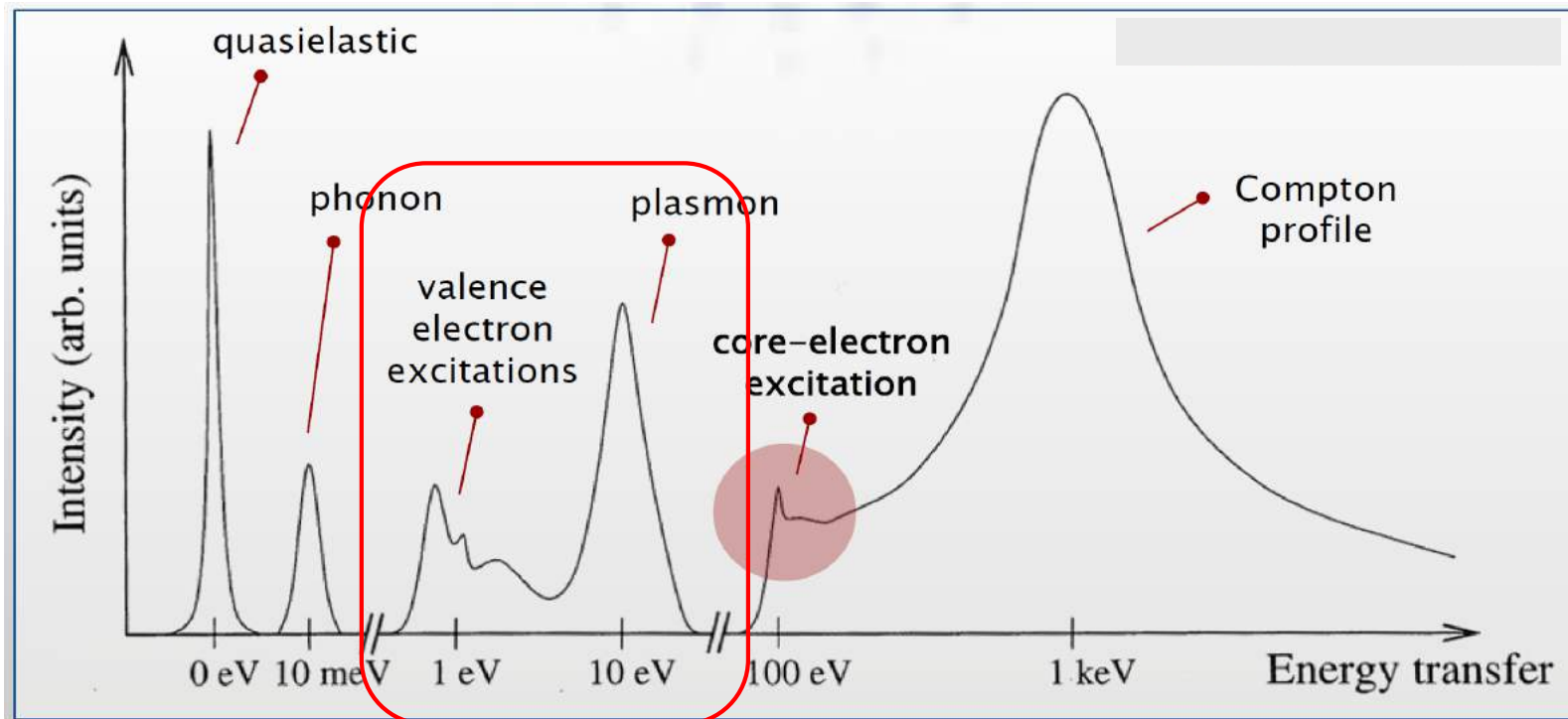
direct photoemission



absorption

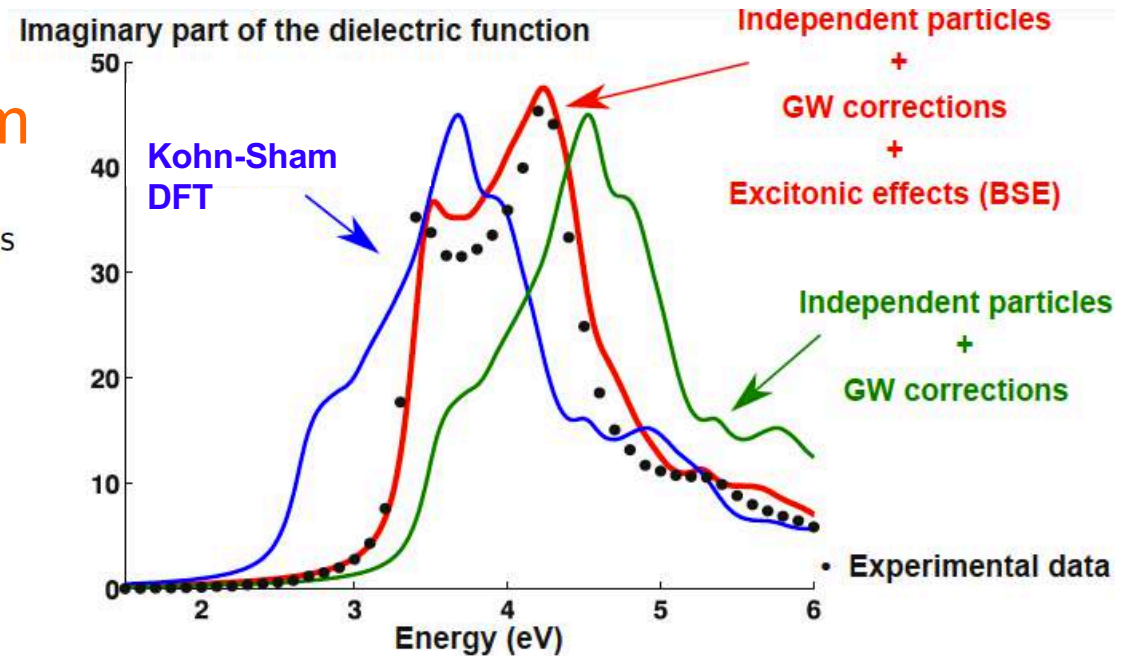
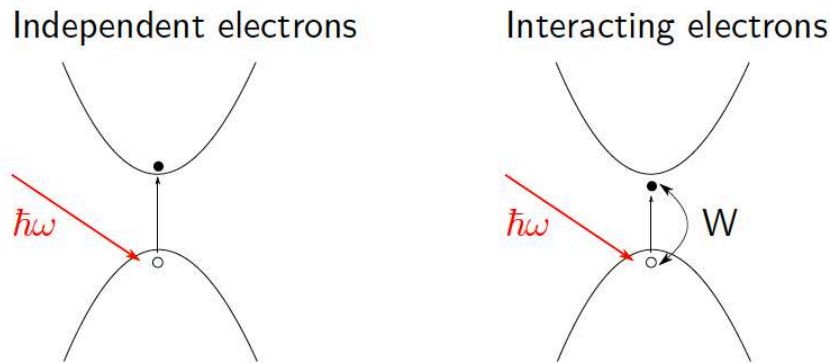


light scattering



# Optical response : absorption

## Silicon absorption spectrum



Beyond **DFT**, with Many-Body Perturbation Theory ... need :

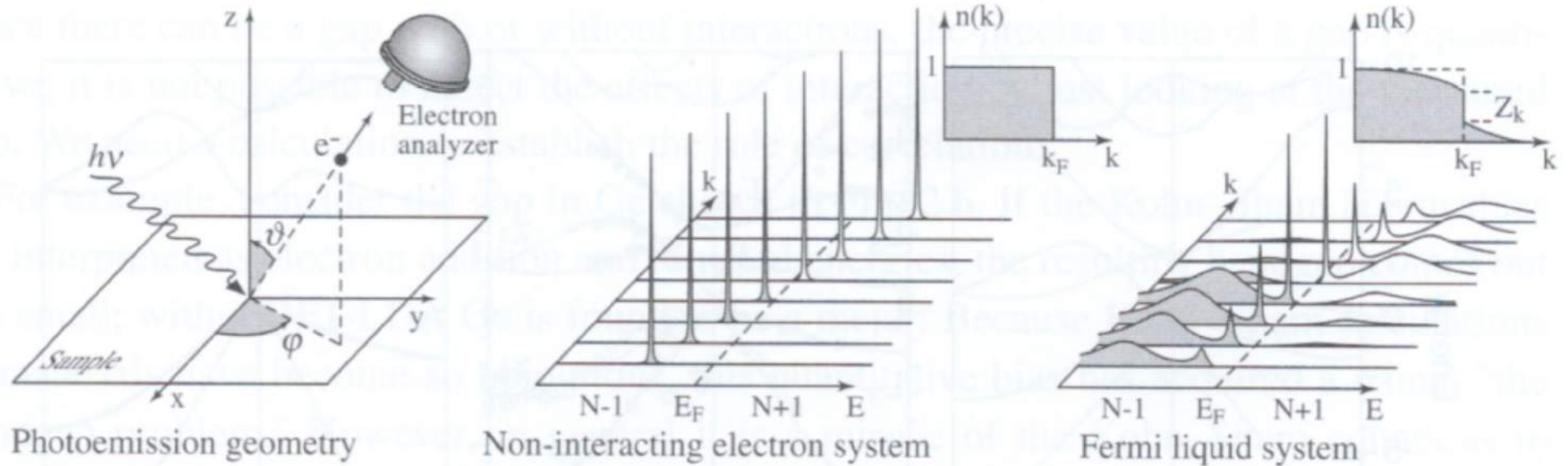
- **band gap correction** (GW) [still independent particle IPA]
- **excitonic corrections** (Bethe-Salpeter Equation - BSE)

Optical spectrum, relation with the dielectric function

$$\text{Im } \epsilon(\omega) = 1 + 4\pi \text{Im } \chi(\omega) \quad \text{Neutral electronic excitations}$$

See lecture of 23 April

# Photoemission experiment

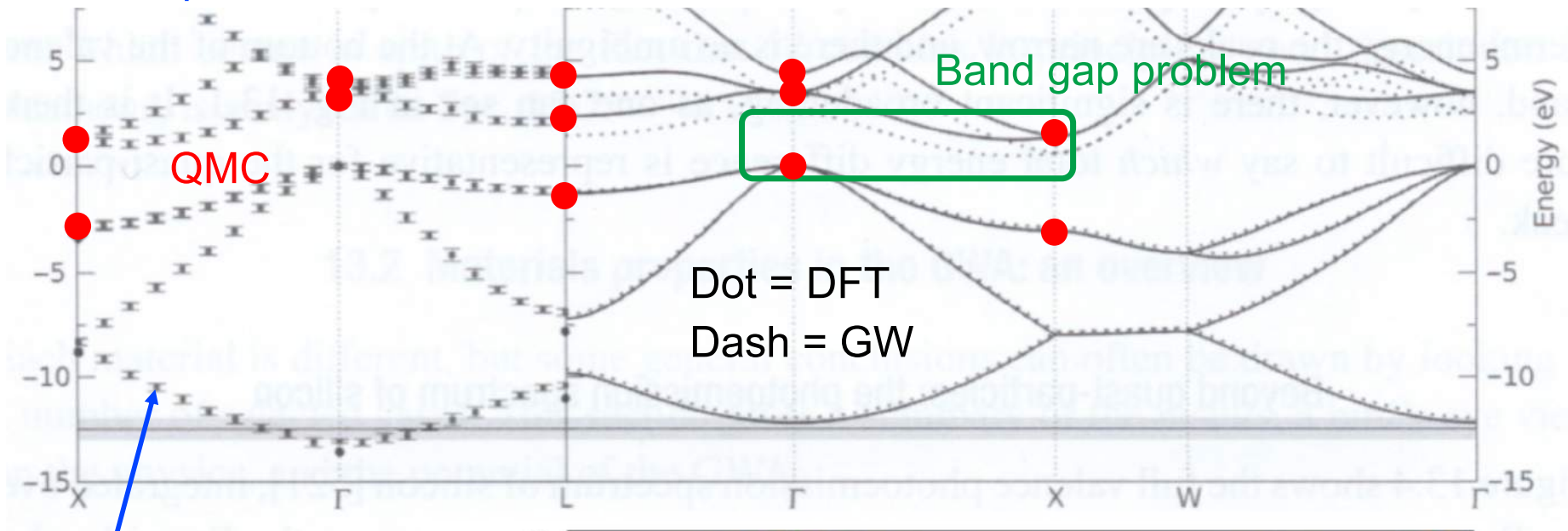


Similar for inverse photoemission ...

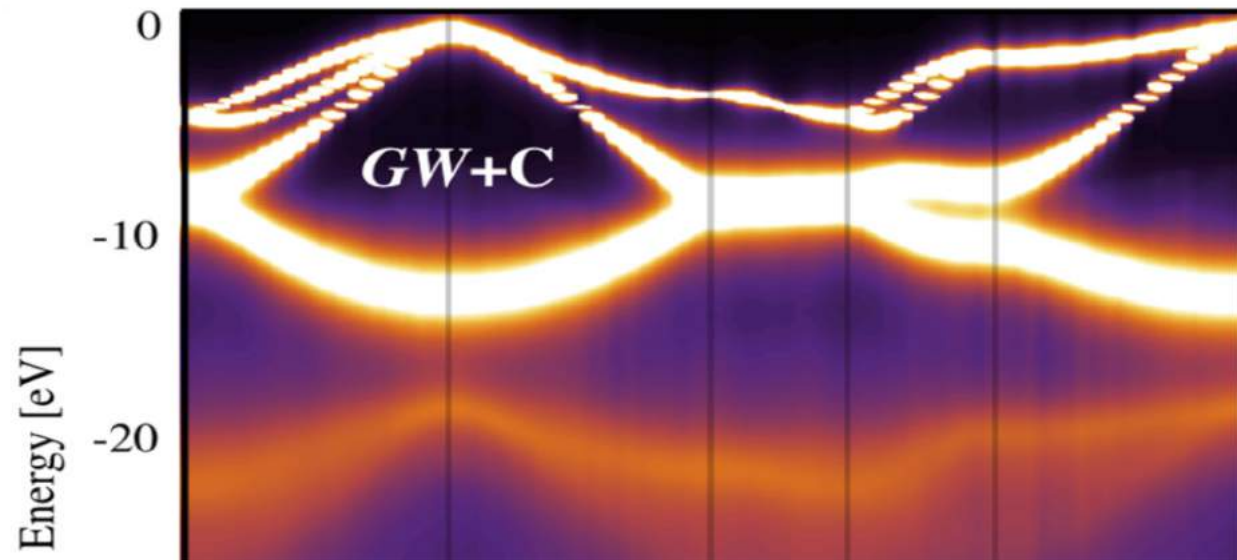
# Silicon band structure

Experiment QMC

Theory DFT GW QMC

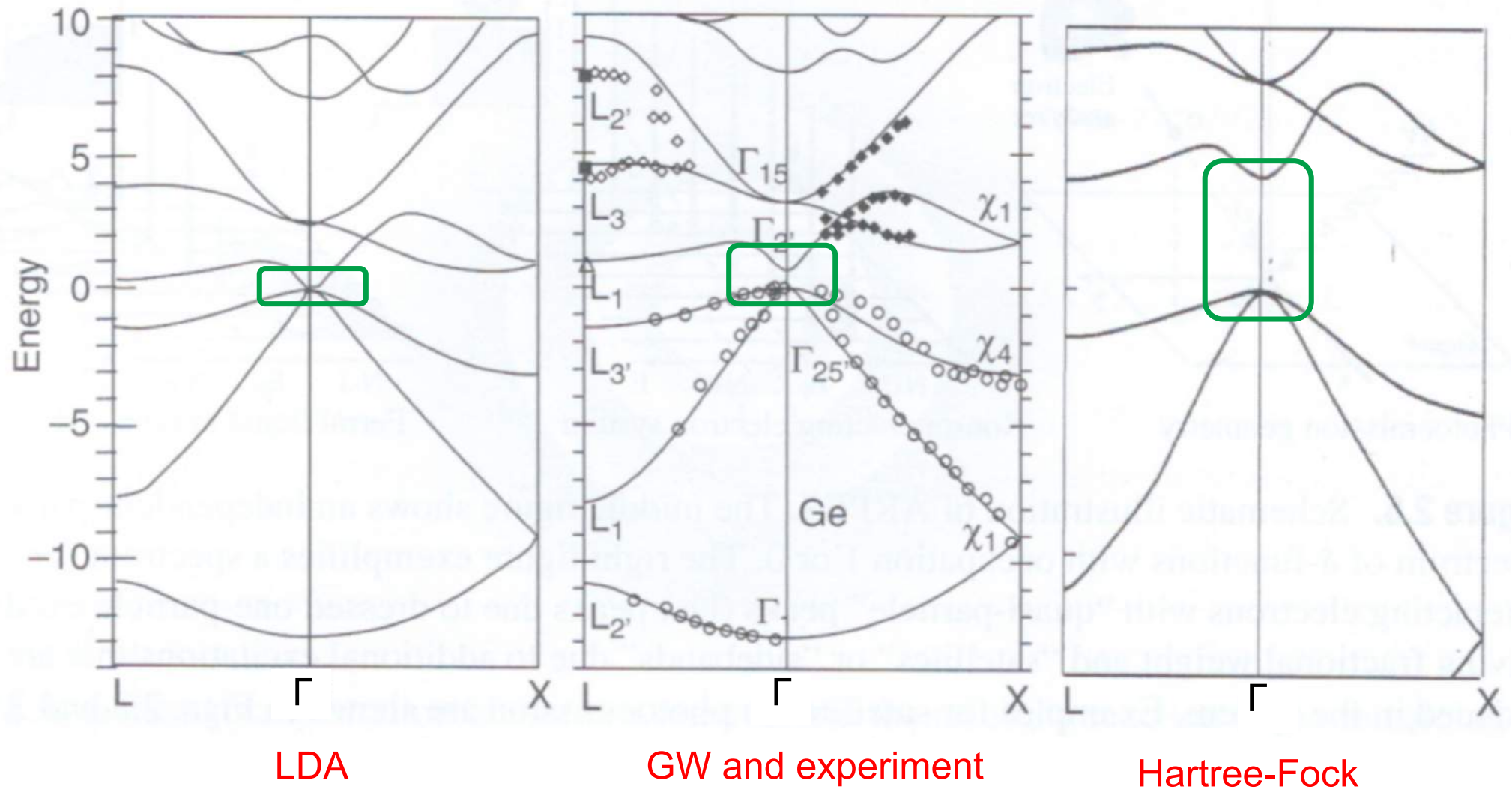


Experimental peaks



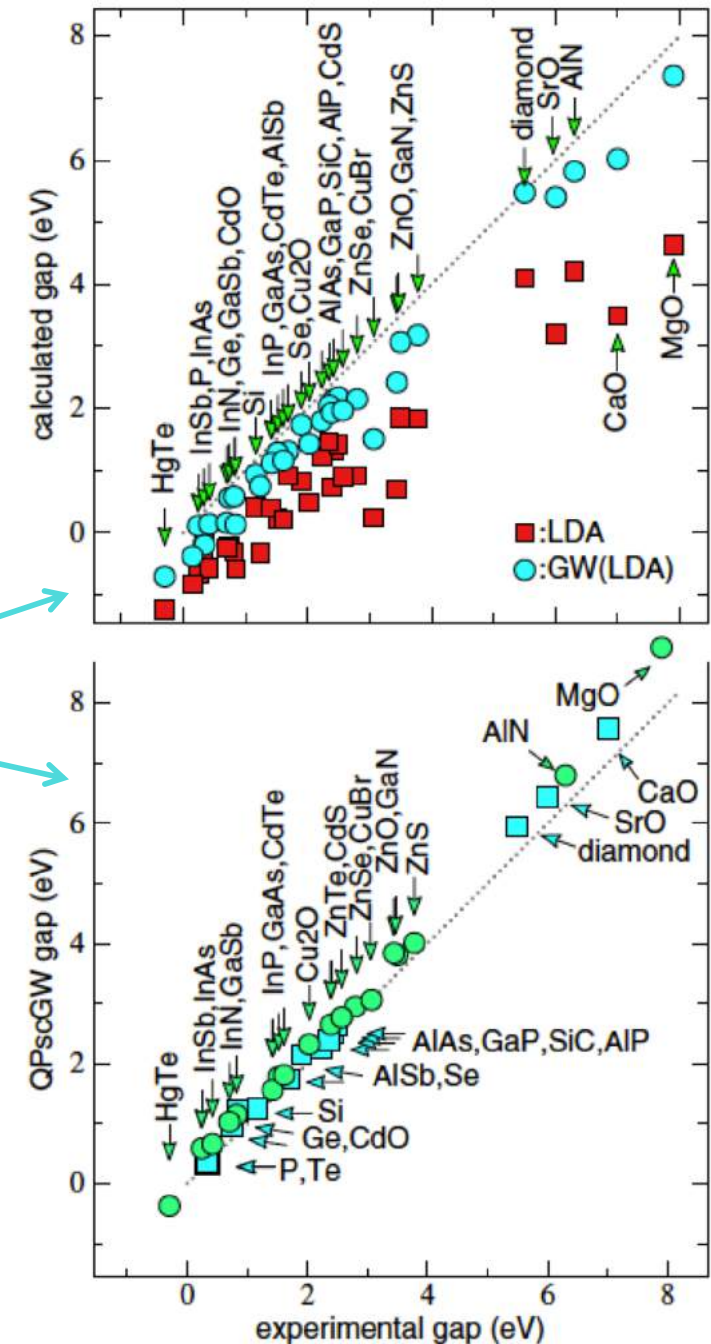


# Germanium band structure



# GW approximation

- Heavier than **DFT/LDA**  
(e.g. CPU time x100...x1000)
- Start from **DFT** calculation,  
then can be one-shot (**GW@LDA**)  
or self-consistent (**SCGW**)
- **SCGW** “Gold standard”  
(within 0.1...0.3eV),  
but **GW@LDA** is often fine



# Charged excitations and Green's function



# Back to undergrad Quantum Mechanics

Starting point : Time-Dependent Schrödinger equation

$$i\hbar \frac{\partial \psi}{\partial t} = \hat{H} \psi$$

e.g. for 1 particle

$$i\hbar \frac{\partial \psi}{\partial t} \Big|_{(\mathbf{r},t)} = -\frac{\hbar^2}{2m} \vec{\nabla}^2 \psi \Big|_{(\mathbf{r},t)} + V(\mathbf{r},t) \psi(\mathbf{r},t)$$

If Hamiltonian does not depend on time,

$$\hat{H} = \sum_{i=1,n} -\frac{\hbar^2}{2m} \nabla_{\mathbf{r}_i}^2 + V(\mathbf{r}_1, \dots, \mathbf{r}_n)$$

then separation of variables

$$\psi(\mathbf{r}_1, \dots, \mathbf{r}_n, t) = \varphi(\mathbf{r}_1, \dots, \mathbf{r}_n) \cdot f(t)$$

$$\Rightarrow i\hbar \frac{\partial f}{\partial t} = E f \quad = \varphi(\mathbf{r}_1, \dots, \mathbf{r}_n) \cdot e^{-i \frac{Et}{\hbar}}$$

Energy : linked to phase acquired upon propagation

# The 1-particle Green's function

Exact ground state wavefunction for  $N$  electrons  $|N,0\rangle$

Creation, annihilation operators  $\psi^+(\mathbf{r},t), \psi(\mathbf{r},t)$

$\psi^+(\mathbf{r},t)|N,0\rangle$  is a  $(N+1)$  electron wavefunction not necessarily in the ground-state (neither in a pure excited state)

$\psi^+(\mathbf{r}',t')|N,0\rangle$  is another  $(N+1)$  electron wavefunction

Let's compare them !

# The 1-particle Green's function

$$iG^e(\mathbf{r}t, \mathbf{r}'t') = \langle N, 0 | \psi(\mathbf{r}, t) \psi^\dagger(\mathbf{r}', t') | N, 0 \rangle \quad \text{for} \quad t > t'$$

Measures how an **extra electron** propagates from  $\mathbf{r}'t'$  to  $\mathbf{r}t$

$$iG^h(\mathbf{r}'t', \mathbf{r}t) = \langle N, 0 | \psi^\dagger(\mathbf{r}', t') \psi(\mathbf{r}, t) | N, 0 \rangle \quad \text{for} \quad t' > t$$

Measures how an **extra hole** propagates from  $\mathbf{r}t$  to  $\mathbf{r}'t'$

$$G(\mathbf{r}t, \mathbf{r}'t') = G^e(\mathbf{r}t, \mathbf{r}'t') - G^h(\mathbf{r}'t', \mathbf{r}t)$$

Compact expression that describes propagation of  
either an extra electron or an extra hole,  
depending on time-ordering

# Lehmann representation

$$iG^e(\mathbf{r}t, \mathbf{r}'t') = \langle N, 0 | \psi(\mathbf{r}, t) \psi^+(\mathbf{r}', t') | N, 0 \rangle$$

Closure relation based on  
complete set of eigenvectors  
with eigenenergies  $E(N \pm 1, i)$

$$\sum_{M,i} |M, i\rangle \langle M, i|$$

$$iG^h(\mathbf{r}'t', \mathbf{r}t) = \langle N, 0 | \psi^+(\mathbf{r}', t') \psi(\mathbf{r}, t) | N, 0 \rangle$$

Then, Fourier transform,

$$G(\mathbf{r}, \mathbf{r}'; \omega) = \sum_i \frac{f_i(\mathbf{r}) f_i^*(\mathbf{r}')}{\omega - \varepsilon_i \pm i\eta}$$

where

$$\varepsilon_i = \begin{cases} E(N+1, i) - E(N, 0) \\ E(N, 0) - E(N-1, i) \end{cases}$$

are **exact excitation energies** (affinities or ionisation) !

# GW approximation to the electronic band structure

# Many-body perturbation theory

Historically older than DFT (from the 40-50' s) !

Big names : Feynman, Schwinger, Hubbard, Hedin, Lundqvist

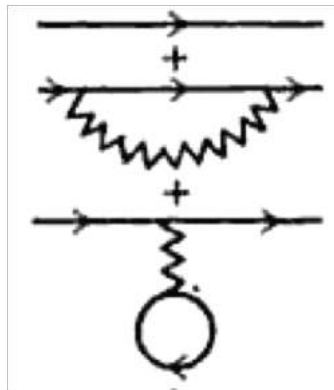
Heavy formalism ! Not amenable to 1 hour presentation ...

Diagrammatic representation of perturbative terms

Based on Green's functions = propagators

1-particle

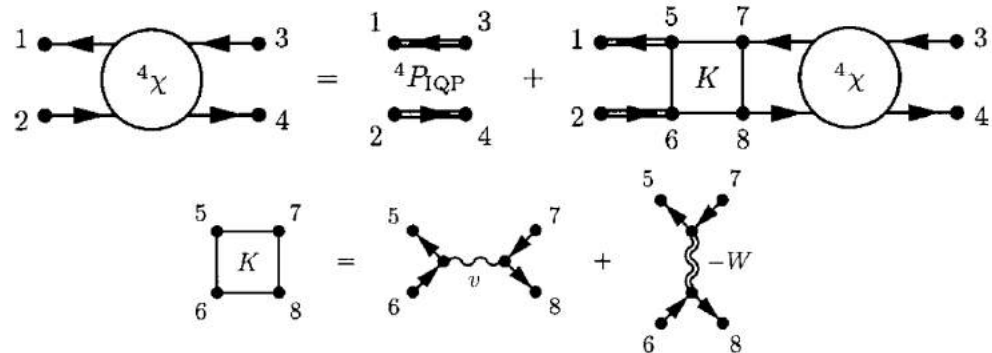
$$G(\mathbf{r}, t; \mathbf{r}', t')$$



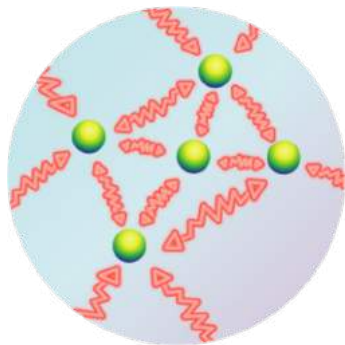
Propagation without Coulomb interactions  
+  
Propagation with polarization of medium  
+  
Propagation with mean-field electrostatic interaction with other electrons

2-particles

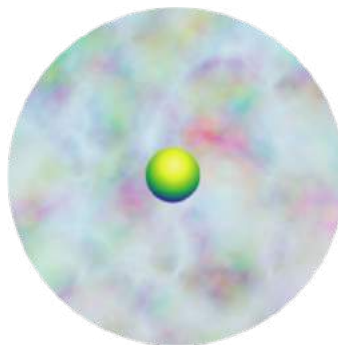
$$G_2(\mathbf{r}_1, t_1; \mathbf{r}_2, t_2; \mathbf{r}_3, t_3; \mathbf{r}_4, t_4)$$



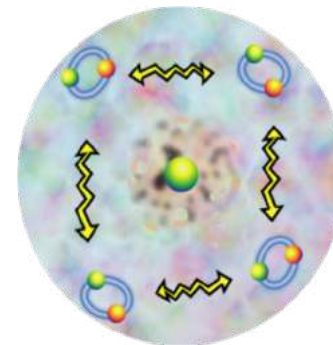
# GW approximation



Wavefunctions



DFT



GW

$$\left( -\frac{1}{2}\nabla^2 + V_{ext}(\mathbf{r}) + V_H(\mathbf{r}) \right) \psi_i(\mathbf{r}) + V_{xc}(\mathbf{r})\psi_i(\mathbf{r}) = \varepsilon_i^{KS} \psi_i(\mathbf{r}) \quad \text{DFT}$$

$$\left( -\frac{1}{2}\nabla^2 + V_{ext}(\mathbf{r}) + V_H(\mathbf{r}) \right) \psi_i(\mathbf{r}) + \int \Sigma(\mathbf{r}, \mathbf{r}'; \varepsilon_i^{QP}) \psi_i(\mathbf{r}') d\mathbf{r}' = \varepsilon_i^{QP} \psi_i(\mathbf{r})$$

$$\Sigma(\mathbf{r}, \mathbf{r}'; \omega) = \lim_{\delta \rightarrow 0^+} \frac{i}{2\pi} \int d\omega' e^{i\omega'\delta} G(\mathbf{r}, \mathbf{r}'; \omega + \omega') W(\mathbf{r}, \mathbf{r}'; \omega') \quad \text{GW}$$

Self energy

Green's function

Screened interaction

# Screened interaction and susceptibility

Computation of the **susceptibility** (for screening of Coulomb)

Adler-Wiser expression, see 23 April

$$\chi_0(\mathbf{r}, \mathbf{r}'; \omega) = - \sum_{\substack{n \\ occ}} \sum_{\substack{m \\ unocc}}^{\infty} \left( \frac{\varphi_n^*(\mathbf{r}) \varphi_m(\mathbf{r}) \varphi_m^*(\mathbf{r}') \varphi_n(\mathbf{r}')}{\varepsilon_m - \varepsilon_n - \omega - i\delta} + \frac{\varphi_n(\mathbf{r}) \varphi_m^*(\mathbf{r}) \varphi_m(\mathbf{r}') \varphi_n^*(\mathbf{r}')}{\varepsilon_m - \varepsilon_n + \omega + i\delta} \right)$$

Approximation : from independent-particle approximation  
Usually starting from KS eigenenergies and eigenfunction

Double sum over a large number of independent « transitions »

$$\varepsilon(\mathbf{r}, \mathbf{r}'; \omega) = \delta(\mathbf{r} - \mathbf{r}') - \int V_C(\mathbf{r} - \mathbf{r}'') P(\mathbf{r}'', \mathbf{r}'; \omega) d\mathbf{r}''$$

Dielectric function

$$W(\mathbf{r}, \mathbf{r}'; \omega) = \int \varepsilon^{-1}(\mathbf{r}, \mathbf{r}''; \omega) V_C(\mathbf{r}'' - \mathbf{r}) d\mathbf{r}''$$

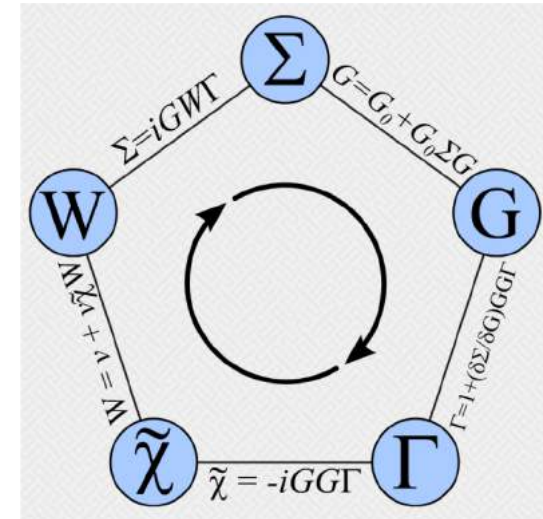
Screened interaction

$$G(\mathbf{r}, \mathbf{r}'; \omega) = \sum_n \frac{\varphi_n(\mathbf{r}) \varphi_n^*(\mathbf{r}')}{\omega - \varepsilon_n}$$

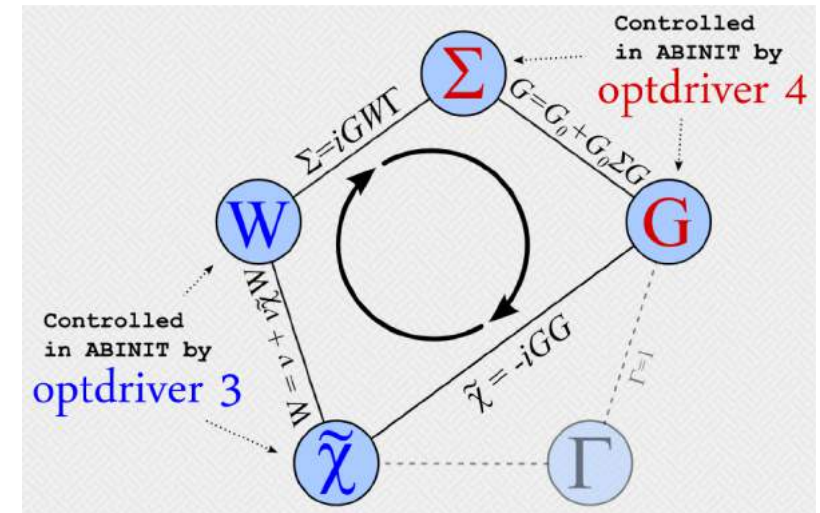
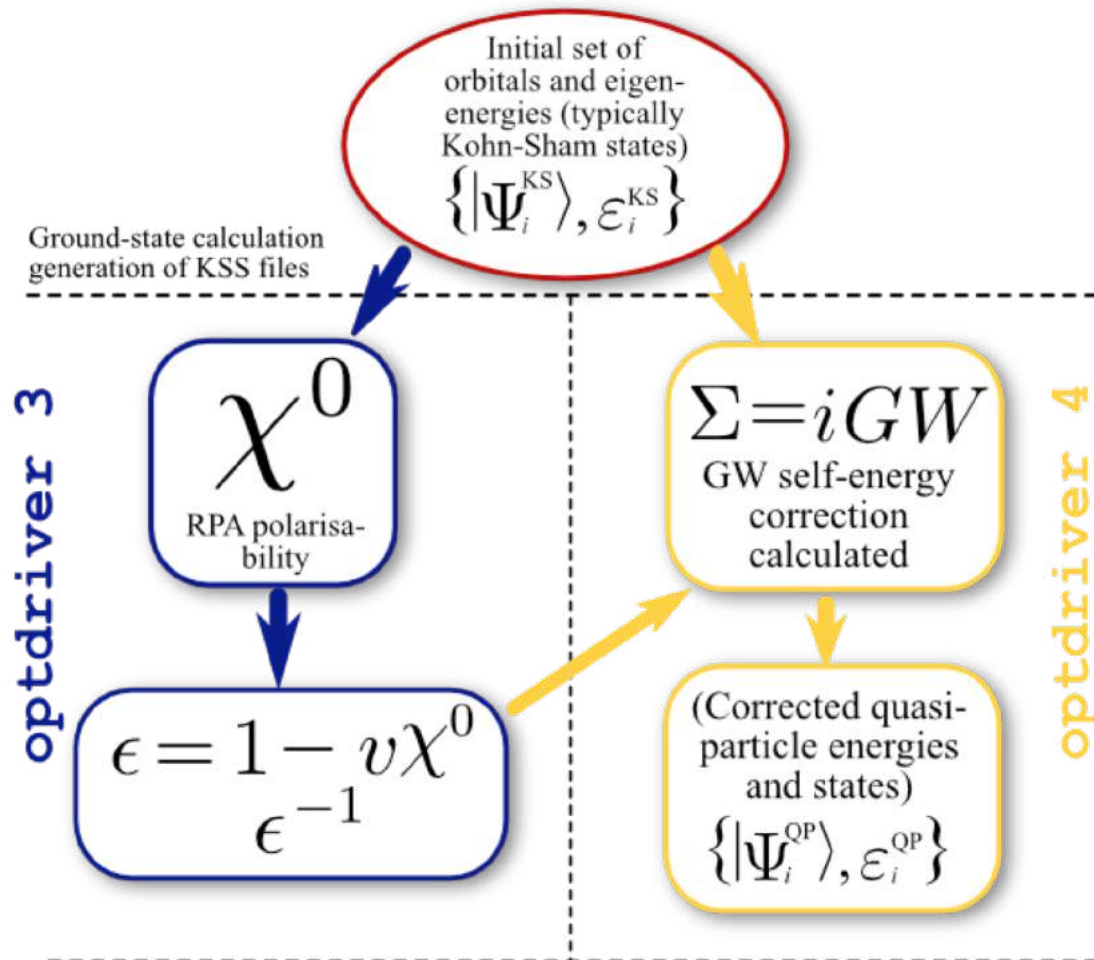
Usually starting from  
KS eigenenergies and eigenfunction



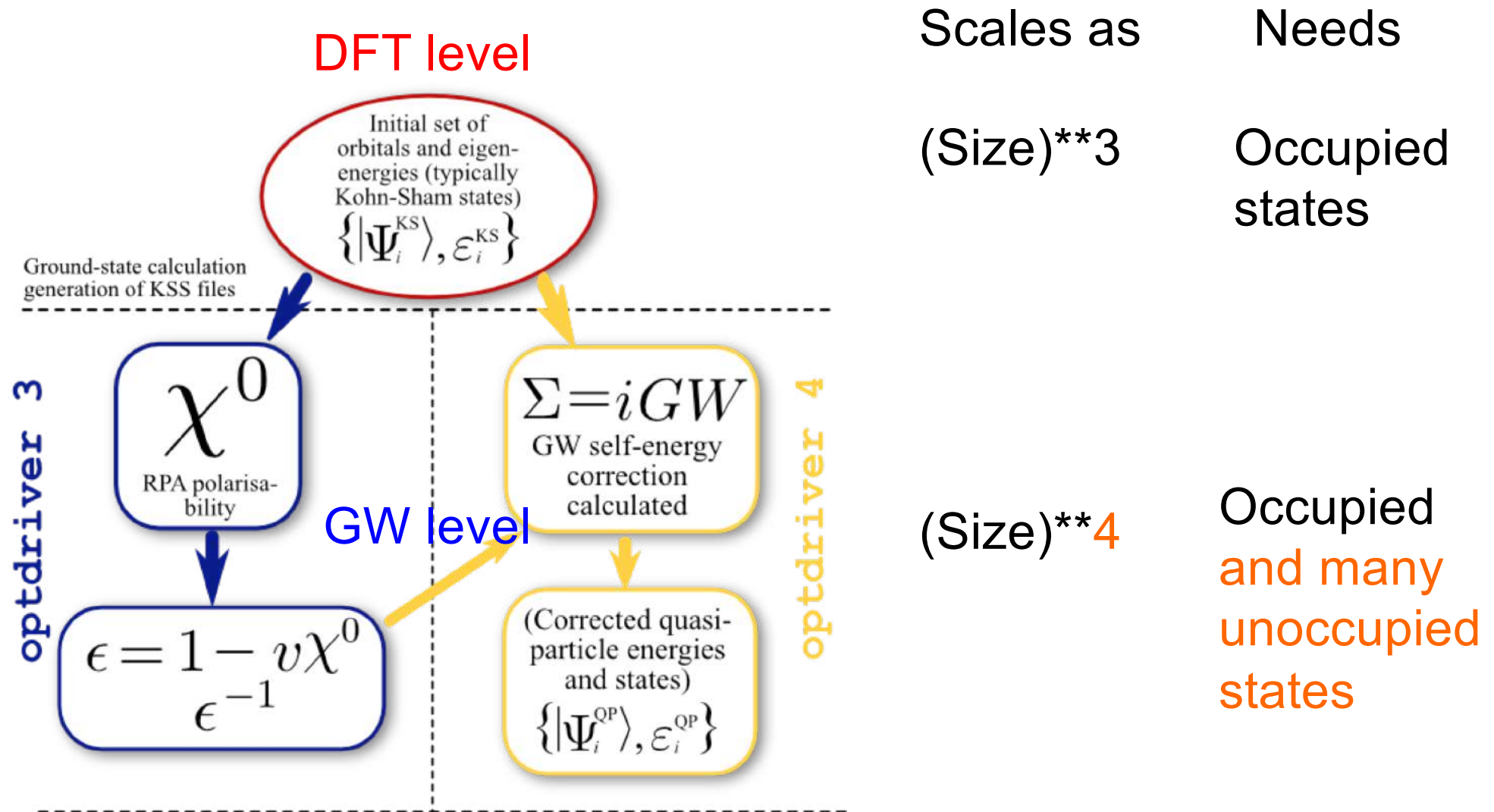
# Self-consistent GW



Neglect of  $\Gamma$   
(vertex corrections)



# GW computational bottlenecks



Scales as

Needs

$(\text{Size})^{**3}$

Occupied states

$(\text{Size})^{**4}$

Occupied and many unoccupied states

DFT : routinely 200 atoms ; GW : routinely 20 atoms

# Avoiding full self-consistency

$G_0W_0$  : DFT  $\Rightarrow G_0, W_0 \Rightarrow \Sigma \Rightarrow$  compute eigenenergies STOP

$$\left( -\frac{1}{2} \nabla^2 + V_{ext}(\mathbf{r}) + V_H(\mathbf{r}) \right) \psi_i(\mathbf{r}) + \int \Sigma(\mathbf{r}, \mathbf{r}'; \epsilon_i^{QP}) \psi_i(\mathbf{r}') d\mathbf{r}' = \epsilon_i^{QP} \psi_i(\mathbf{r})$$

$GW_0$  : DFT  $\Rightarrow G_0, W_0 \Rightarrow \Sigma \Rightarrow$  eigenenergies+eigenfunctions  
then iterate on  $G \Rightarrow \Sigma \Rightarrow$  compute eigenenergies

$$G(\mathbf{r}, \mathbf{r}'; \omega) = \sum_n \frac{\varphi_n(\mathbf{r}) \varphi_n^*(\mathbf{r}')}{\omega - \epsilon_n}$$

$e$ - $GW$  : DFT  $\Rightarrow G_0, W_0 \Rightarrow \Sigma \Rightarrow$  compute eigenenergies  
then update energies in  $G, W \Rightarrow \Sigma \Rightarrow$  compute eigenenergies

$$\chi_0(\mathbf{r}, \mathbf{r}'; \omega) = - \sum_{\substack{n \\ occ}} \sum_{\substack{m \\ unocc}}^{\infty} \left( \frac{\varphi_n^*(\mathbf{r}) \varphi_m(\mathbf{r}) \varphi_m^*(\mathbf{r}') \varphi_n(\mathbf{r}')}{\epsilon_m - \epsilon_n - \omega - i\delta} + \frac{\varphi_n(\mathbf{r}) \varphi_m^*(\mathbf{r}) \varphi_m(\mathbf{r}') \varphi_n^*(\mathbf{r}')}{\epsilon_m - \epsilon_n + \omega + i\delta} \right)$$

# Plasmon-pole models to decrease CPU time

$$\chi_0(\mathbf{r}, \mathbf{r}'; \omega) = - \sum_{\substack{n \\ occ}} \sum_{\substack{m \\ unocc}}^{\infty} \left( \frac{\varphi_n^*(\mathbf{r}) \varphi_m(\mathbf{r}) \varphi_m^*(\mathbf{r}') \varphi_n(\mathbf{r}')}{\epsilon_m - \epsilon_n - \omega - i\delta} + \frac{\varphi_n(\mathbf{r}) \varphi_m^*(\mathbf{r}) \varphi_m(\mathbf{r}') \varphi_n^*(\mathbf{r}')}{\epsilon_m - \epsilon_n + \omega + i\delta} \right)$$

Scales as (Size)\*\*4

Sparing evaluations at many frequencies ?

Plasmon-pole models (PPM) allows one to deduce frequency dependence from either static susceptibility only, or static and dynamic susceptibility at only one frequency.  
⇒ one order of magnitude speed-up.

Most often : Godby-Needs plasmon pole

# Linearized quasi-particle equation

$$\left( -\frac{1}{2}\nabla^2 + V_{ext}(\mathbf{r}) + V_H(\mathbf{r}) \right) \psi_i^{KS}(\mathbf{r}) + V_{xc}(\mathbf{r}) \psi_i^{KS}(\mathbf{r}) = \epsilon_i^{KS} \psi_i^{KS}(\mathbf{r})$$

$$\left( -\frac{1}{2}\nabla^2 + V_{ext}(\mathbf{r}) + V_H(\mathbf{r}) \right) \psi_i^{QP}(\mathbf{r}) + \int \Sigma(\mathbf{r}, \mathbf{r}'; \epsilon_i^{QP}) \psi_i^{QP}(\mathbf{r}') d\mathbf{r}' = \epsilon_i^{QP} \psi_i^{QP}(\mathbf{r})$$

Self-energy  $\Sigma(\vec{r}, \vec{r}'; \epsilon_i^{QP})$  depends on  $\epsilon_i^{QP}$  ?!

Rewrite quasi-particle equation :

$$\left( -\frac{1}{2}\nabla^2 + V_{ext}(\mathbf{r}) + V_H(\mathbf{r}) + V_{xc}(\mathbf{r}) \right) \psi_i^{QP}(\vec{r}) + \int \left( \Sigma(\mathbf{r}, \mathbf{r}'; \epsilon_i^{QP}) - V_{xc}(\mathbf{r}) \right) \psi_i^{QP}(\mathbf{r}') d\mathbf{r}' = \epsilon_i^{QP} \psi_i^{QP}(\mathbf{r})$$

Treat difference between self-energy and  $V_{xc}$  as perturbation

$$\epsilon_i^{QP} = \epsilon_i^{KS} + \left\langle \psi_i^{KS} \left| \Sigma(\epsilon_i^{QP}) - V_{xc} \right| \psi_i^{KS} \right\rangle$$

Moreover 
$$\Sigma(\epsilon_i^{QP}) = \Sigma(\epsilon_i^{KS}) + (\epsilon_i^{QP} - \epsilon_i^{KS}) \left. \frac{\partial \Sigma}{\partial \epsilon} \right|_{\epsilon_i^{KS}}$$

Thus 
$$\epsilon_i^{QP} = \epsilon_i^{KS} + Z_i \left\langle \psi_i^{KS} \left| \Sigma(\epsilon_i^{KS}) - V_{xc} \right| \psi_i^{KS} \right\rangle$$

with 
$$Z_i^{-1} = 1 - \left\langle \psi_i^{KS} \left| \left. \frac{\partial \Sigma}{\partial \epsilon} \right|_{\epsilon_i^{KS}} \right| \psi_i^{KS} \right\rangle$$

# Further topics ...

Implementation using plane waves and pseudopotentials / PAW

Convergence parameters

Specific treatments for 2-dimensional and 1-dimensional systems

Scaling : algorithmic improvements