Accuracy and Validation of Results

Georg KRESSE

Institut für Materialphysik and Center for Computational Material Science

Universität Wien, Sensengasse 8 4, A-1090 Wien, Austria

UNIVERSITÄT WIEN

Overview

- How is the precision controlled in VASP
  - the plane wave energy cutoff
  - technical errors
  - the critical parameters `ENAUG, ENCUT, LREAL, ROPT`
  - the “super-flag” `PREC`
- Related issues
  - k-point sampling
  - slab thickness

Energy cutoff

- controls the completeness of the basis set

  at each k-point only the plane waves that fulfil

$$\frac{\hbar^2}{2m_e}|\mathbf{G}+\mathbf{k}|^2 < E_{\text{cutoff}}$$

  are included

  different number of plane waves at each k-point

- E_{cutoff} is controlled by `ENCUT` in the INCAR file

  the number of plane wave for each k-point is written to the OUTCAR file:

```
k-point   1 :   0.25000.25000.2500  plane waves:   1546
k-point   2 :   -.25000.25000.2500  plane waves:   1557
```

- defaults for `ENCUT` are supplied in the pseudopotential files (POTCAR)

  usually the maximum `ENMAX` is chose as energy cutoff

Convergence correction

- VASP applies an automatic convergence correction based on the kinetic energy of wavefunctions in the atomic limit

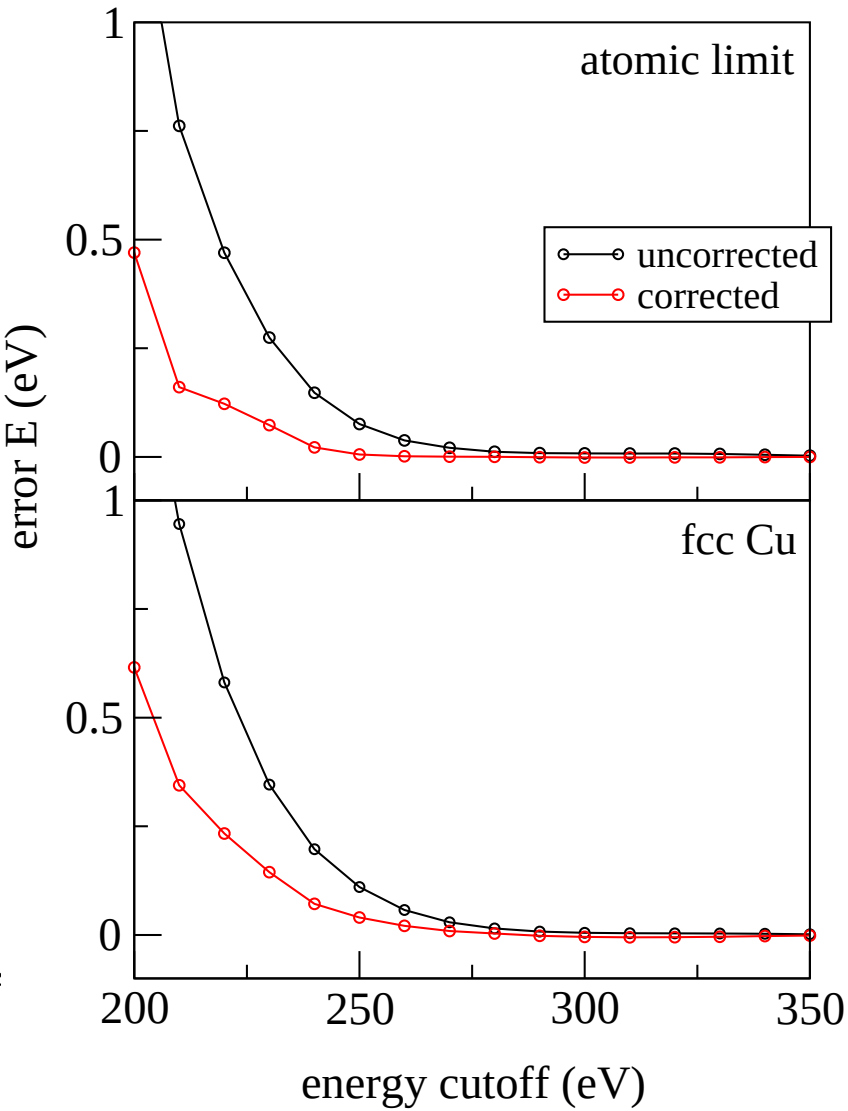
  ```
  energy of atom 1      EATOM=-1393.0707
  kinetic energy error for atom= 0.0229
  ```

- works well in the atomic limit, and for free electron metals

  corrects for 80 % of the total error

- for d-elements and bulk calculations, corrections are only partial

- due to correction, the energy might increase when the cutoff is increased



Can you rely on the default cutoff ?

it depends

- `ENCUT` is a very reasonable compromise between accuracy and speed
- you can rely on `ENCUT`, as long as the

  cell-shape and the volume remain unchanged

  – frozen phonon calculations
  – surface and slab calculations
  – adsorption of molecules on surfaces
- otherwise you might need to be rather careful

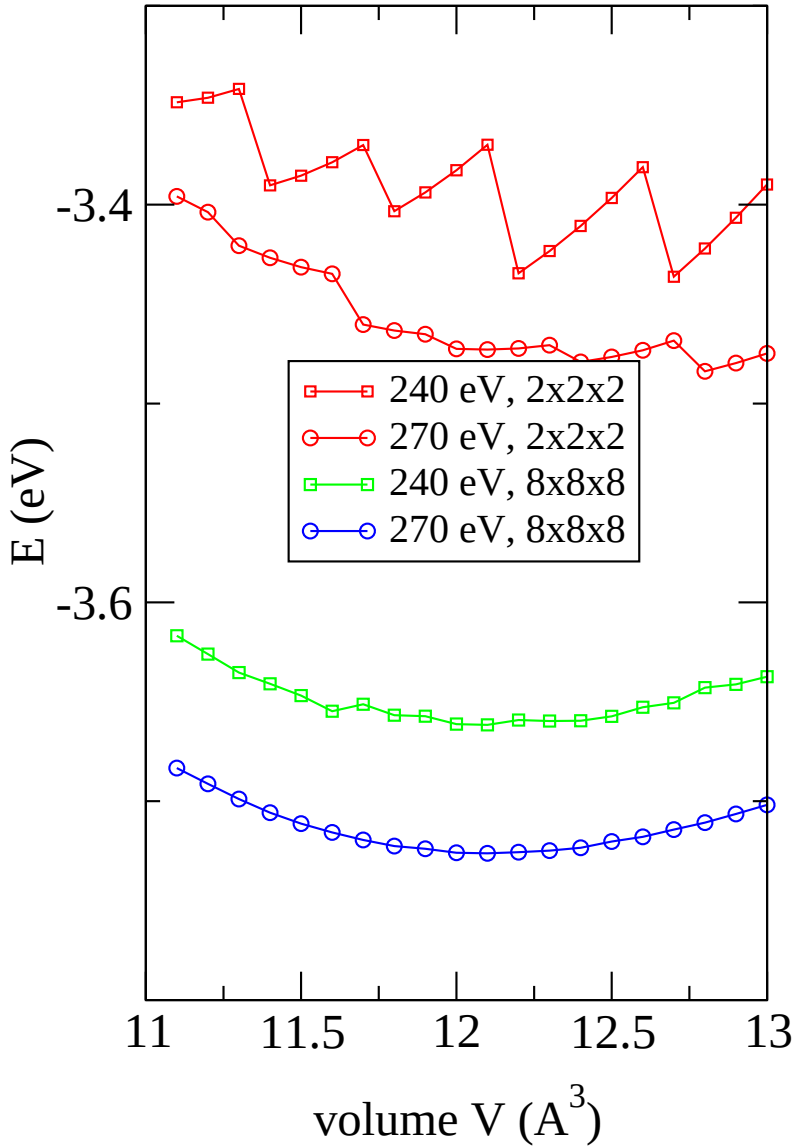
  the basis set changes discontinuously when the cell-shape is changed, since new plane waves are included when they satisfy the cutoff criterion

$$\frac{\hbar^2}{2m_e}|\mathbf{G}+\mathbf{k}|^2 < E_{\text{cutoff}}$$

k-points and cutoff

energy versus volume for fcc Cu

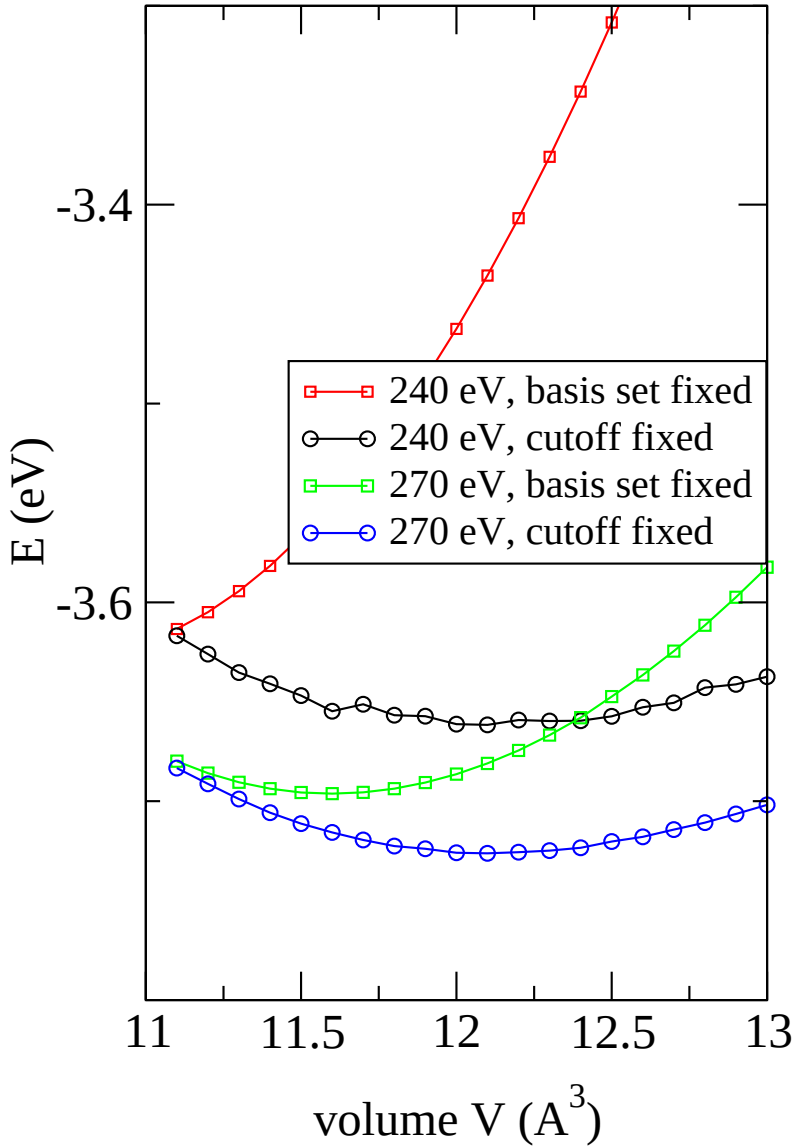
- by using more k-points or a higher energy cutoff, the energy surface becomes smoother
  at 270 eV and using 8x8x8 k-points, the energy veries smoothly
- in general, elastic constants are most prone to such errors
  if you sample the energy surface on a coarse scale, problems are less severe (recommended distortions 1 %)



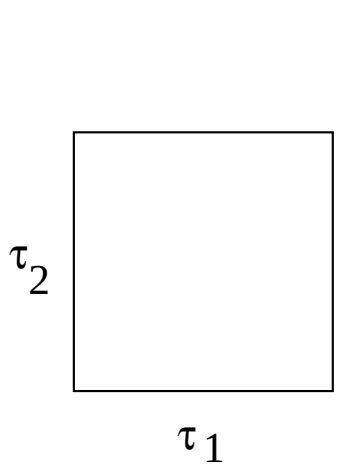
Fixed basis-sets instead of fixed cutoff

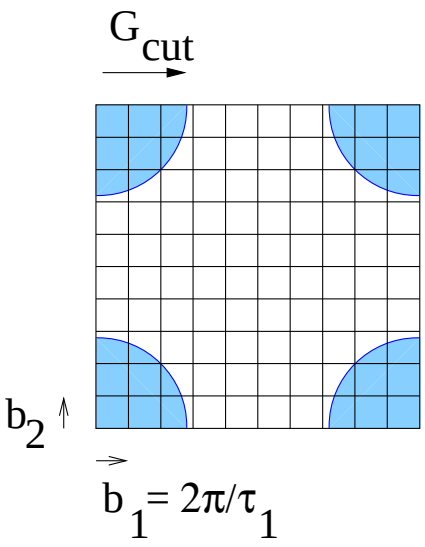
- possible by restarting with `ISTART=2` but such calculations clearly yield much too small volumes even at 270 eV (5 % error)
- effectively the cutoff decreases when the volume is increased (since the reciprocal lattice vectors become shorter)
- fixed basis set calculations are

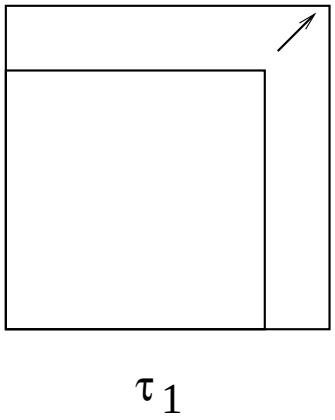
  obviously a very bad idea

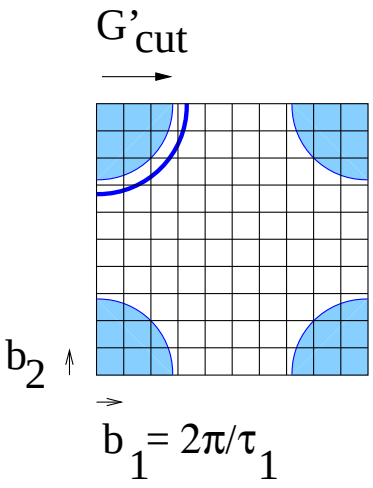


Fixed basis-set calculations





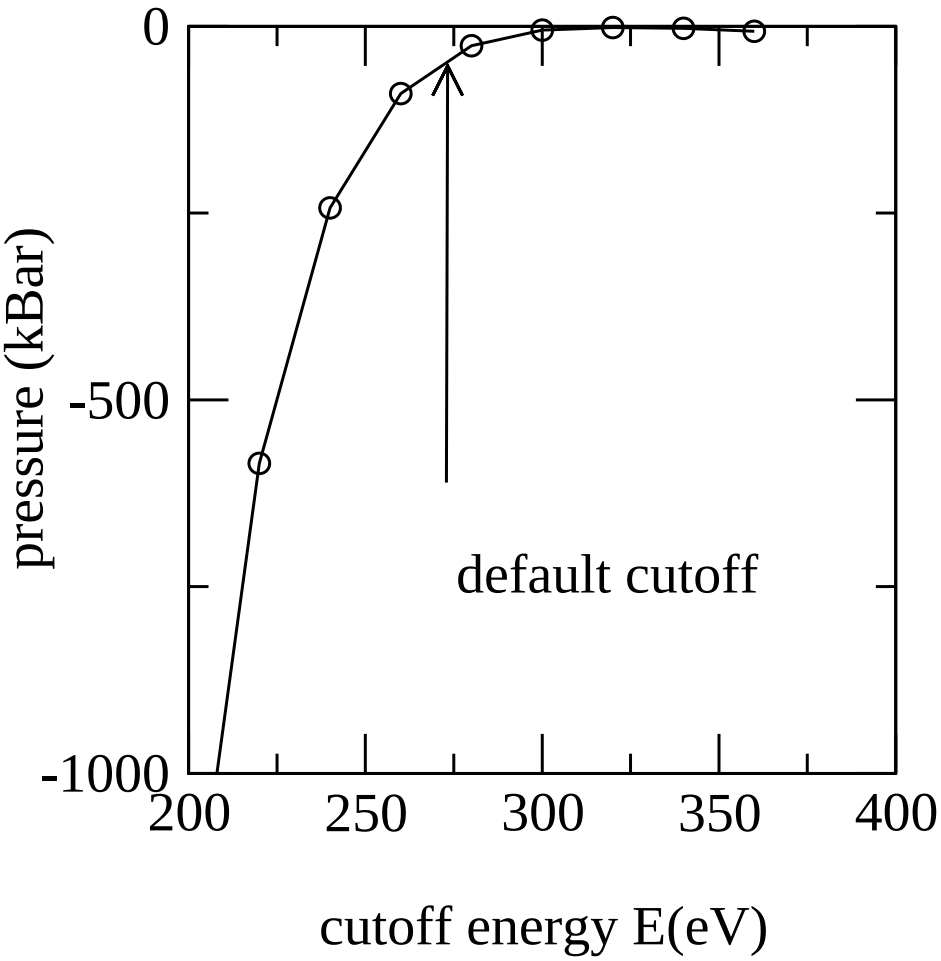




- the cutoff decreases by a factor τ_1/τ_1' when the lattice is expanded from $\tau_1 \rightarrow \tau_1'$
- for the expanded lattice the basis set corresponds effectively to a lower cut-off G_{cut}' and therefore a lower quality,
  $\Rightarrow$ the energy is overestimated at larger volumes
  $\Rightarrow$ the volume is underestimated for fixed basis-set calculations

Stress tensor

- the stress tensor is implicitly calculated at a fixed basis set
  upon cell-shape or volume relaxation one obtains too small volumes
  (2-5 % errors at the default cutoff)
- cutoff must be increased by 20-30%, when cell relaxations are performed
- calculations at the equilibrium lattice parameter of fcc Cu:
  270 eV: p=−50 kBar (contract)
  350 eV: a few kBar (correct result)



Cell-shape relaxations

- increase the cutoff by 30 %

  and restart the calculations, after the first ionic relaxation has succeeded
  the basis set is then adopted to the new geometry

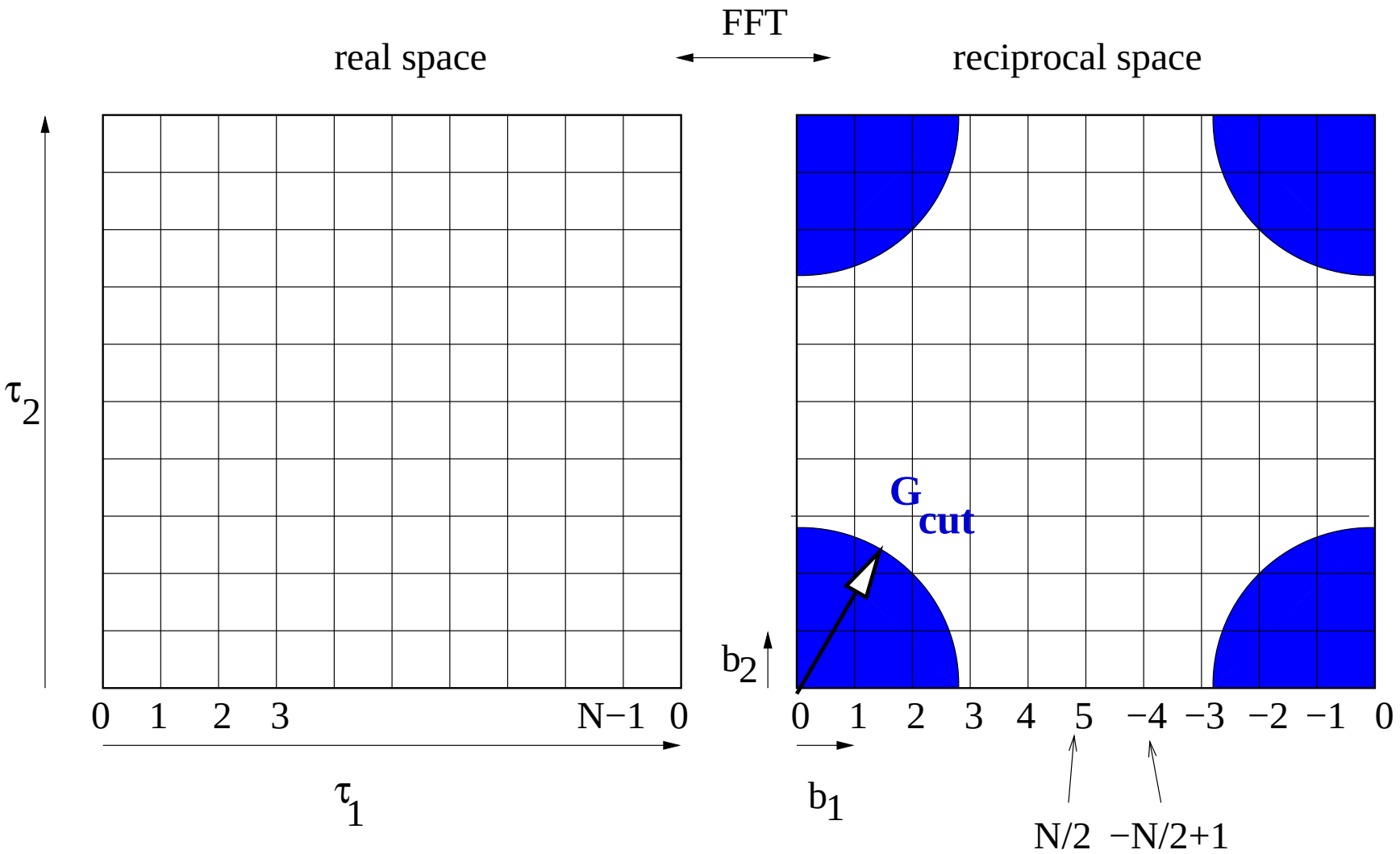
- quick and dirty (if you need to save computer time)

  the error in the stress tensor is rather uniform, and it can be supplied in the INCAR file

  – calculate the stress tensor at a larger energy cutoff
  – calculate the stress tensor at the desired low energy cutoff
  – supply the difference of the average of the diagonal elements of the stress tensor (pressure) in the INCAR file (should be a negative value)

    ```
    PSTRESS = p(low cutoff)-p(high cutoff)
    ```

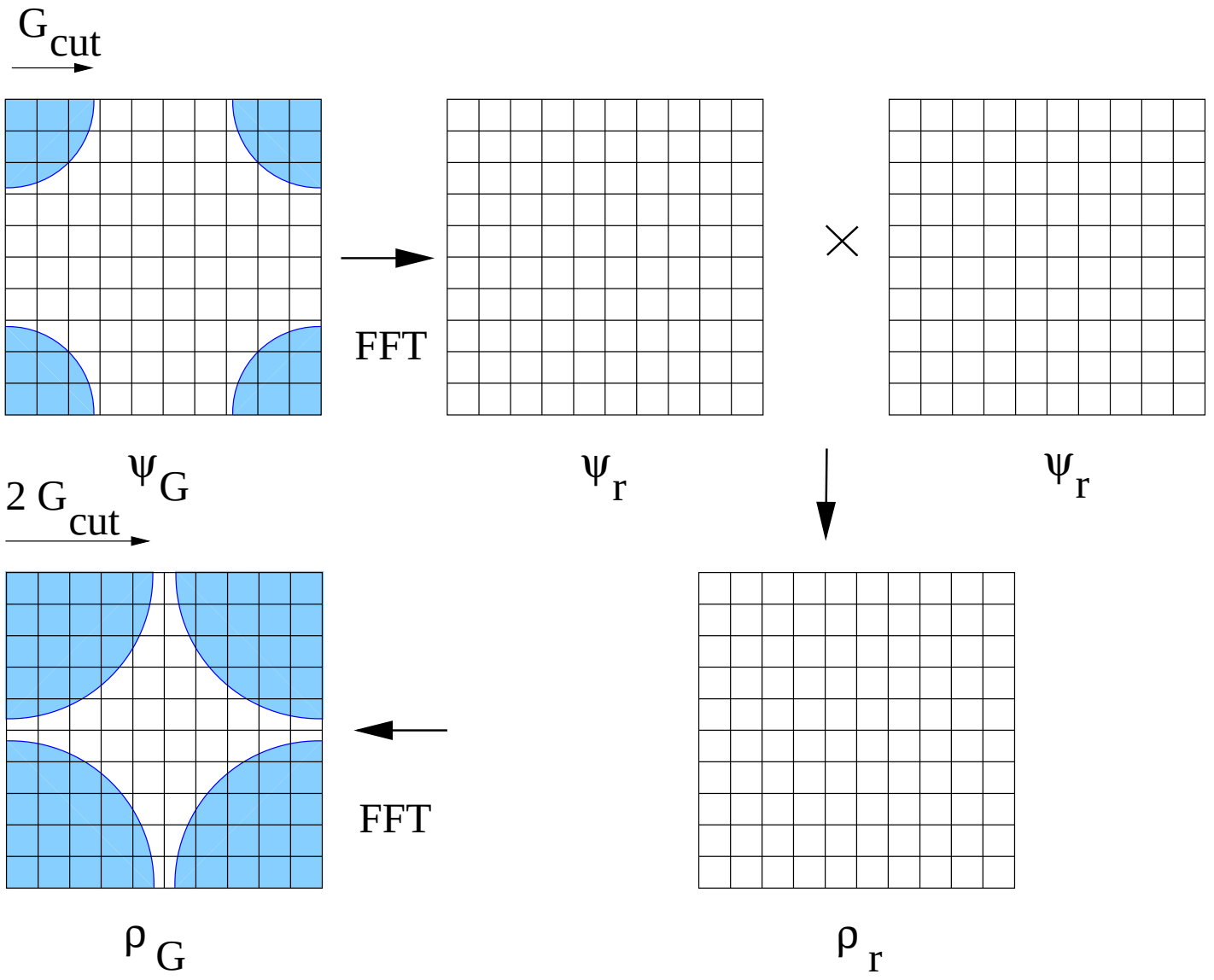
Technical errors related to the truncated FFT mesh



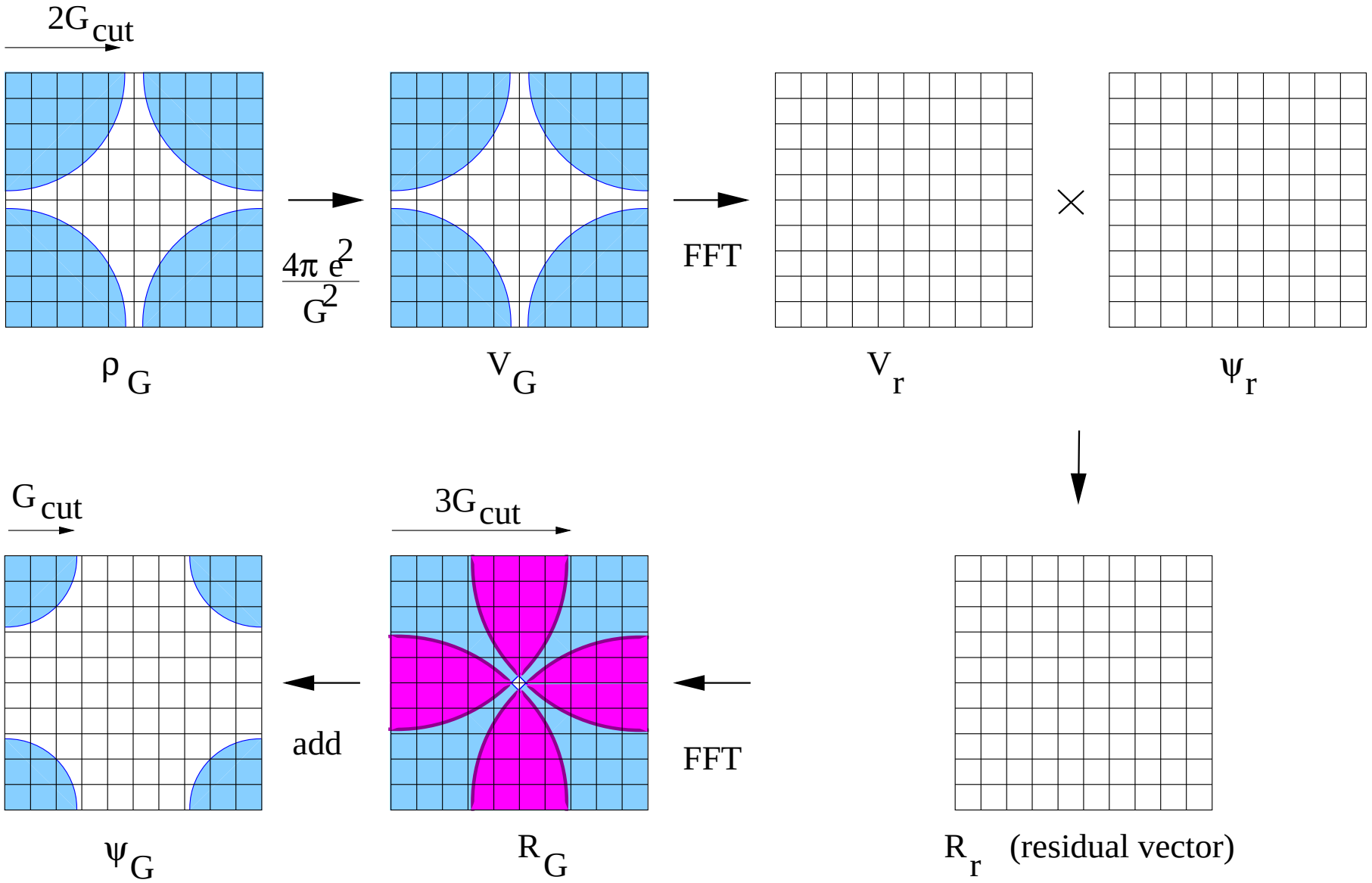
$x_1 = n_1 / N \;\; \tau_1$

$g_1 = n_1 \, 2\pi / \tau_1$

Evaluation of the charge density



Evaluation of the local part of the Hamiltonian **Hψ**



The FFT grid

- the folding theorem implies that the charge density contains components up to $2 \times G_{\text{cut}}$ where

$$\frac{\vec{\hbar}^2}{2m_e}|G_{\text{cut}}|^2 = E_{\text{cut}}$$

- the Hartree potential contains Fourier components up to $2 \times G_{\text{cut}}$ as well
- the residual vector contains Fourier components up to $3 \times G_{\text{cut}}$

to avoid any errors, the Fourier grid must contain all wave-vectors up to $2 \times G_{\text{cut}}$

this is true for both, the evaluation of the charge-density and the residual vector

- if this is not the case, components in the charge density are wrapped around from the other side of the box: “wrap around errors”

  the proper terminus technicus is aliasing errors

  high frequency components are aliased to low frequency components
  (similar to AD converters, where you perform oversampling to avoid such errors)

What sort of errors does this cause

- the translational invariance is destroyed

  if all atoms are shifted by an arbitrary vector $\vec{\tau}$ the energy should remain exactly identical

  this is however only the case, if aliasing errors are avoided

  equivalently, the sum of all ionic forces should be zero

$$\sum_{i=1}^{N_{\text{atoms}}} \vec{F}_i = 0$$

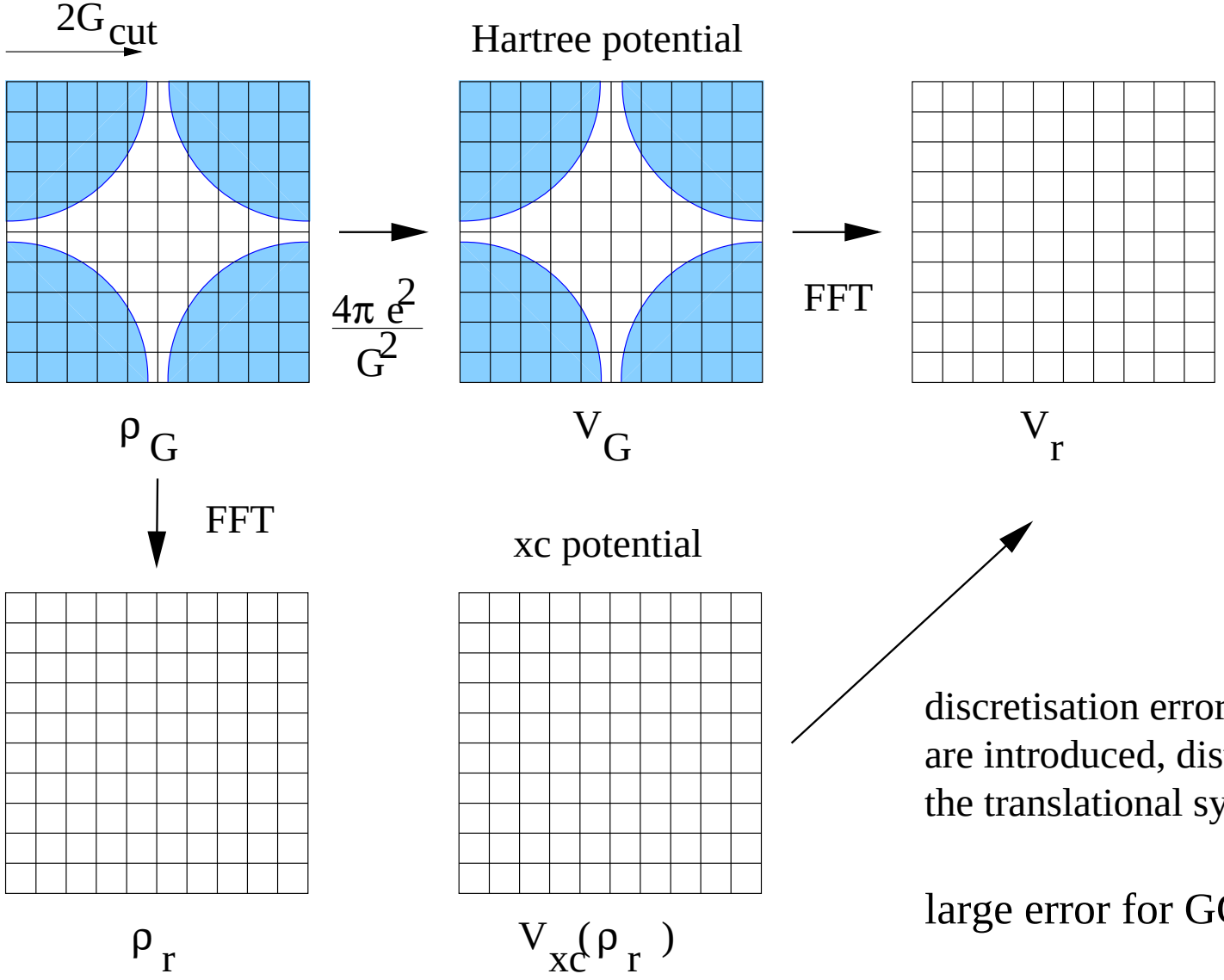
  offers a convenient way to check for such errors

- symmetry inequivalent atoms are no longer strictly symmetry equivalent

  VASP however symmetrises the charge and the forces explicitly

  to quantify this sort of errors, you need to switch off symmetry `ISYM=0`

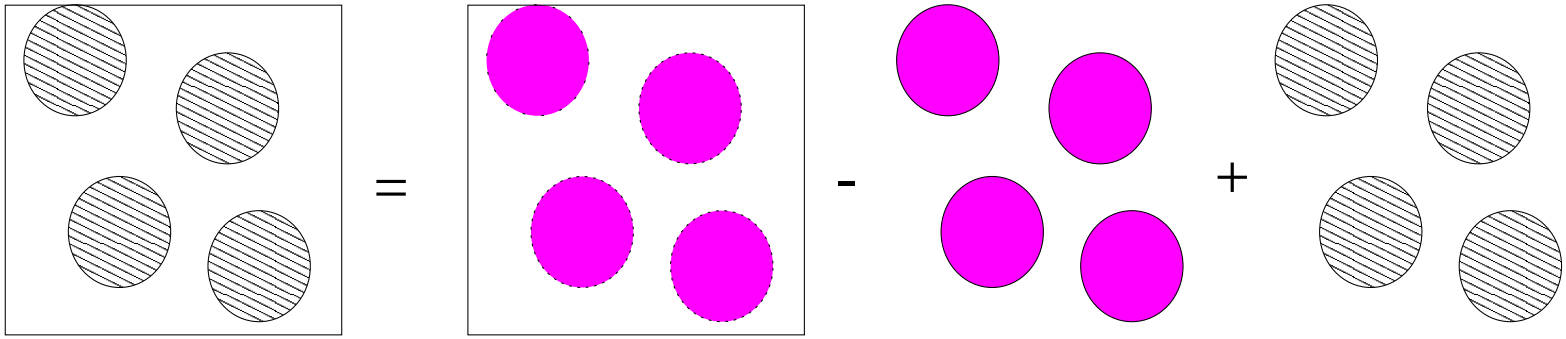
Exchange correlation potential



The PAW compensation charge on regular grid

- the pseudo-wavefunctions do not have the same norm as the AE wavefunctions inside the spheres
- to deal with long range electrostatic interactions between spheres

  a "soft" compensation charge $\hat{n}$ is introd. (similar to FLAPW)

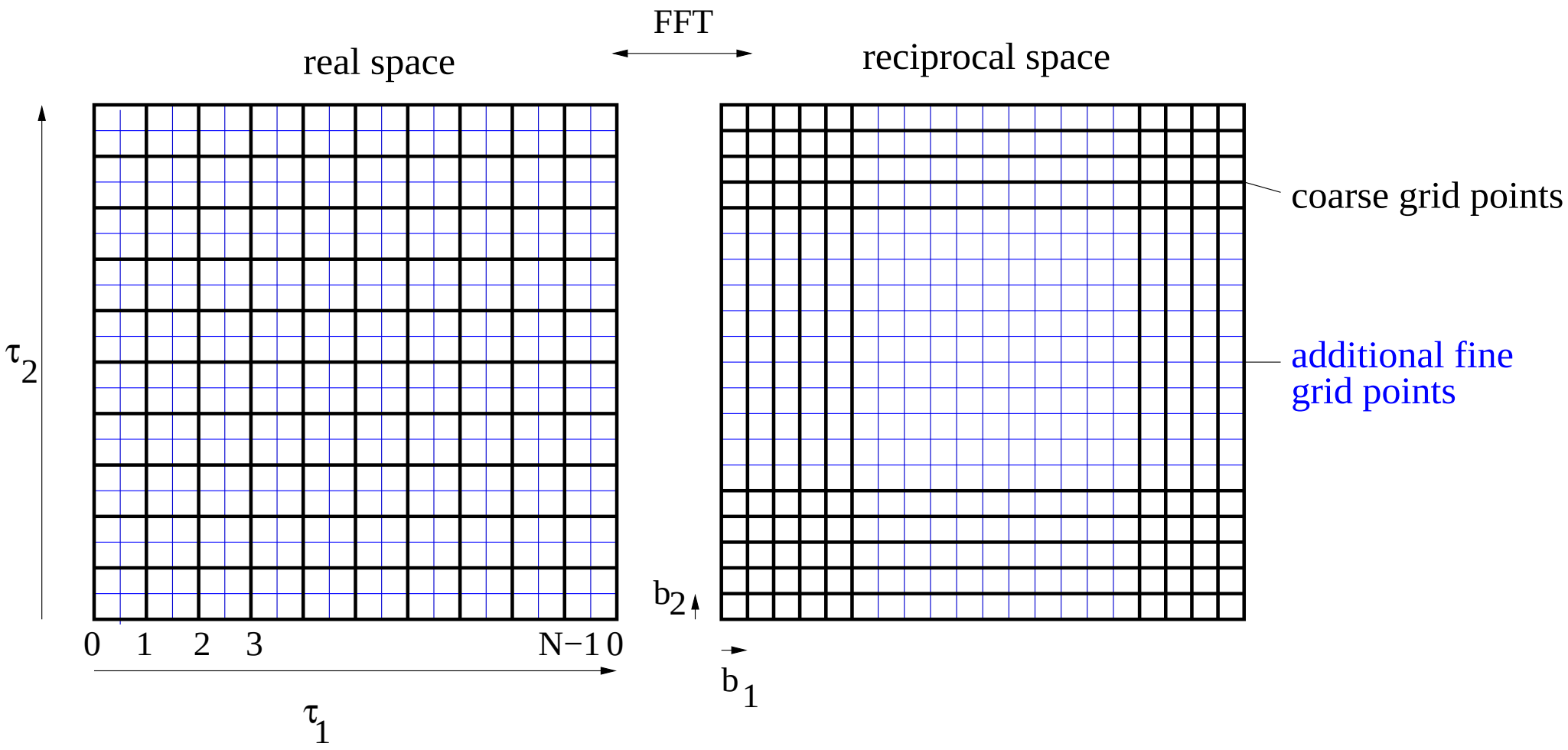


AE pseudo + compens. pseudo+comp. onsite AE-onsite

- these localised compensation charges can be rather hard and are not well represented on the plane wave grid

  $\Rightarrow$ dual grid technique

Representation of the compensation charge: Dual grid technique



data transfer occurs only in reciprocal space, grids are not necessarily “aligned”

evaluation of the potentials (XC) is also done on the fine grid

Controlling the aliasing errors in VASP

- in VASP, the coarse (plane wave) FFT grid is controlled by the INCAR parameters `NGX`, `NGY` and `NGZ`
- for the default setting (`PREC=Medium`, or `PREC=Normal`) VASP sets `NGX`, `NGY` and `NGZ` such that all wave vectors up to $3/2 \times G_{\text{cut}}$ are included

  this causes a small wrap around or "aliasing" error
- in VASP, the second (finer) FFT grid is controlled by the INCAR parameters `NGXF`, `NGYF` and `NGZF`
- Jürgen Furthmüllers `fftlib` supports only radices of 2, 3, 5 and 7 and the FFT dimensions must be dividable by 2

  i.e. only certain values are allowed for `NGx` and `NGxF`

$$22^{n_2} 3^{n_3} 7^{n_5} 5^{n_7}$$

Non local part of the potentials

- in the
  PAW and PP methods, the following expressions occurs in the evaluation of the $\mathbf{H}|\psi_n\rangle$

  $$\sum_{ij}^{\text{sites}} |\tilde{p}_j\rangle D_{ij} \langle \tilde{p}_i|\psi_n\rangle$$

- the expression can be evaluated in real space or reciprocal space

$$C_{in} = \langle \beta_i|\psi_n\rangle = \frac{\Omega}{N_{\text{FFT}}} \sum_{\mathbf{r}} \langle \beta_i|\mathbf{r}\rangle\langle \mathbf{r}|\psi_{n\mathbf{k}}\rangle = \frac{\Omega}{N_{\text{FFT}}} \sum_{\mathbf{r}} \beta_i(\mathbf{r})\psi_{n\mathbf{k}}(\mathbf{r}) =$$

$$= \sum_{\mathbf{G}} \langle \beta_i|\mathbf{k}+\mathbf{G}\rangle\langle \mathbf{k}+\mathbf{G}|\psi_{n\mathbf{k}}\rangle = \sum_{\mathbf{G}} \bar{\beta}(\mathbf{k}+\mathbf{G}) C_{\mathbf{G}n}.$$

- in reciprocal space $N_{\text{planewaves}} \times N_{\text{ion}} \times N_{\text{proj}}$ operations are required
  $\mathbf{H}|\psi_n\rangle$ scales quadratically with the number of ions
- in real space $N_{\text{points}} \times N_{\text{ion}} \times N_{\text{proj}}$ operations are required, since $\beta_i(\mathbf{r})$ is localised around ions $\Rightarrow \mathbf{H}|\psi_n\rangle$ scales linearly with the number of ions

Aliasing errors due to real space projection

- the projector function $\beta_i(\mathbf{r})$ must be optimised in order to remove all high frequency components, without affecting their accuracy

  high frequency components are experienced as noise in the calculations (again translational symmetry is removed)

- the most recent version of the real space projection scheme should be selected by specifying `LREAL = A`utomatic in the INCAR file

  Optimization of the real space projectors (new method)

- real space optimisation has also side effects

  the absolute energies are slightly modified, and hence calculations with and without real space optimisation should not be compared

- the real space optimisation is controlled by the `ROPT` parameter

  ```
  ROPT = -1E-2 to -2E-4   (smaller abs. value is better)
  ```

  one value for each atomic species (each POTCAR file)

Always check the OUTCAR file when LREAL *is used*

```
maximal supplied QI-value         = 16.25
optimisation between [QCUT,QGAM] = [  8.29, 16.74] = [ 19.24, 78.46] Ry
Optimized for a Real-space Cutoff    1.37 Angstroem

  l    n(q)    QCUT    max X(q) W(low)/X(q) W(high)/X(q)  e(spline)
  2      6     8.288     4.974    0.20E-03    0.59E-03    0.30E-07
  2      6     8.288    13.453    0.16E-02    0.43E-02    0.21E-06
  0      7     8.288    13.269    0.32E-04    0.61E-04    0.13E-07
  0      7     8.288    44.490    0.60E-03    0.17E-03    0.30E-06
  1      6     8.288     5.266    0.44E-03    0.24E-03    0.50E-07
  1      6     8.288     7.318    0.14E-02    0.14E-02    0.22E-06
```

W(low)/X(q) is a measure for the modification of the projector functions compared to LREAL=F

W(high)/X(q) is a measure for the noise in the real space projector functions

both decrease when the absolute value of ROPT is decreased

Three sources of aliasing errors

- coarse grid errors
  - charge density steming from the soft part of the wavefunctions $\tilde{\Psi}$
  - application of the local part of the potential to the wavefunctions $\tilde{\Psi}$
- errors stemming from the representation of the soft compensation charges on the second finer grid

  related errors due to the xc-potential
  can be substantial for GGA's
- errors stemming from the non local part of the pseudopotential, when real space projection is selected

`total drift` in forces as written to the OUTCAR file indicates how accurate the calculations are

The PREC *tag*

the PREC tag allows to control the behaviour of VASP in a convenient manner, by influencing a number of other parameters

```
PREC = Low | Medium | High |          Normal | Accurate
```

- Low: only recommended for quick and dirty calculations

  e.g. initial relaxations with few k-points

- Normal: standard calculations

- Accurate: exceptional high accuracy

- the two older settings Medium and High are no longer recommended, although they are still supported

PREC *and* ENCUT

PREC	ENCUT	NGx	NGxF	ROPT
Low	max(ENMIN)	3/2 G_{cut}	3 G_{aug}	-1E-2
Med	max(ENMAX)	3/2 G_{cut}	4 G_{aug}	-2E-3
High	max(ENMAX)*1.3	2 G_{cut}	16/3 G_{aug}	-4E-4
Normal	max(ENMAX)	3/2 G_{cut}	2 NGx	-5E-4
Accurate	max(ENMAX)	2 G_{cut}	2 NGx	-2.5E-4

$$\frac{\hbar^2}{2m_e}|G_{\text{cut}}|^2 = \texttt{ENCUT} \qquad \frac{\hbar^2}{2m_e}|G_{\text{aug}}|^2 = \texttt{ENAUG}$$

max(ENMAX/ENMIN) corresponds to the maximum ENMAX/ENMIN found in POTCAR

ENAUG defaults to the maximum EAUG found in POTCAR

PREC= Normal *and* Accurate

- for Accurate wrap around errors are avoided, whereas for Normal 3/4 of the required grid dimensions are used

  Normal is an excellent compromise

- the energy cutoff ENCUT should be set manually in any case in the INCAR file

  this makes the calculations more concise and better controlled

  for stress calculations and cell shape deformations, one might need to increase ENCUT from the default value

- the grids for the compensations charges have exactly twice the dimension than those of the coarser grids

  (Hartree and XC potentials are also evaluated on those grids)

PREC= Normal offers a very high accuracy at modest computational costs

PREC= Medium *and* High

- for High wrap around errors are avoided as for Accurate

  for Medium 3/4 of the required grid dimensions are used as for Normal

  for High the energy cutoff is increased

  I now recommended to do this manually in the INCAR file

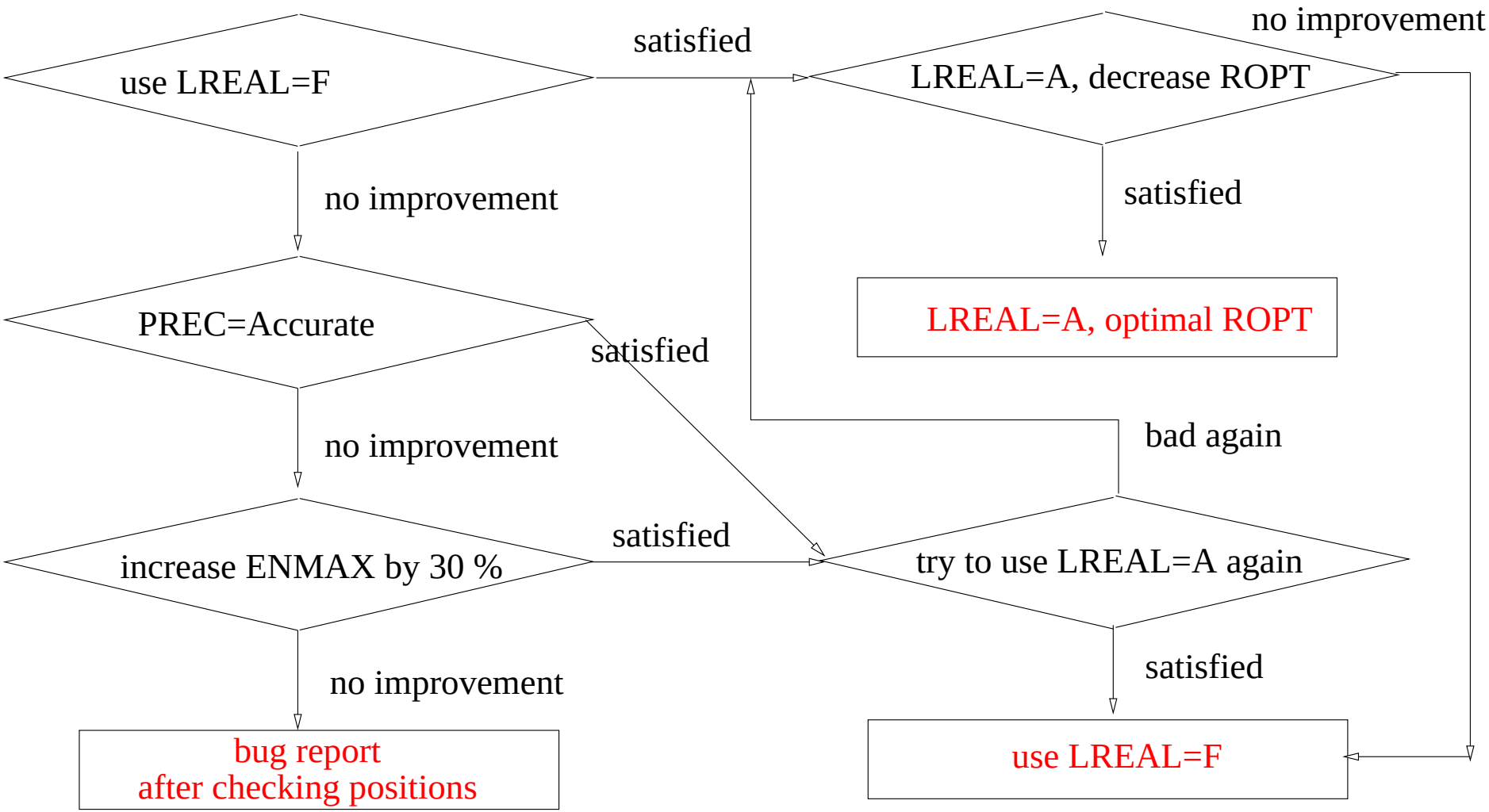
  ENCUT should be specified manually for any calculation

- the defaults for ROPT were not sufficiently accurate for Medium and High

- the grids for the augmentations charges are controlled by ENAUG

  this offers more flexibility, but the "doubled grids" used for Normal and Accurate are more precise and do not cost a noticable amount of computer time

What to do, if you are not satisfied with the forces (drift)



A few points to keep in mind

- the minimal input in the INCAR file is

```
PREC  = Normal  |  Accurate
LREAL = Auto    |  False
ENCUT = xxxx       (ROPT = xxxx  xxxx  xxxx)
```

  calculations done with an identical setup are comparable

- when you use `Medium` or `High`:

```
PREC  = Medium  |  High
LREAL = Auto    |  False
ENCUT = xxxx    ENAUG = xxxx   (ROPT = xxxx  xxxx  xxxx)
```

- never calculate energy difference between calculations with different setups (including k-points)

The most common mistakes

- energy differences from calculations with different energy cutoffs

  Pt slab calculations with 3x3x4 atoms at the default cutoff 230 eV
  add CO molecule and calculate adsorption energy (CO default 400 eV)

  errors will be propotional to the number of Pt atoms and around 200 meV

- energy differences from calculations with different KPOINTS

Validating results

- cutoff and aliasing errors:
  - increase the cutoff
  - or try to perform `PREC=Accurate` calculations
  - possibly switch of the real space optimisation

Related errors:

- increase the number of k-points
- increase the slab thickness
- for defects increase the size of the supercell to remove artificial interactions

TEST, TEST, TEST ....