

The Software Package at Democritos

P. Giannozzi

NEST-INFM, Scuola Normale Superiore di Pisa, giannozz

CINECA, Bologna, 2004/03/01

- Background:
Democritos and scientific software
Why yet another plane-wave-pseudopotential
density-functional package?
On the sad status of scientific software in physics
Free software and GNU license
- Present status:
Technical characteristics: features, algorithms,
coding, installation

Democritos and scientific software

The Democritos center of Italian INFN is dedicated to atomistic simulations of materials, with a strong emphasis on the development of high-quality scientific software.

Democritos is based in Trieste, where a strong tradition of computer simulation of materials (mostly but not exclusively from first principles) exists.

Democritos has an ongoing collaboration¹ with several other institutions (Princeton, UPenn, NY State, UIUC, IBM) aiming at the following goals:

- Development of software for next-generation massively parallel machines ($N > 10000$ processors)
- Production of a high-quality software package for atomistic simulations based on electronic structure, in particular at the density-functional theory (DFT) level, using plane waves (PW) and pseudopotentials (PP).

¹*Open-sesame*: Open-source Scalable Electronic Structure and Atomistic Modeling Environment

Why yet another PW-PP-DFT package?

- because it is not another package, but a merge of several pre-existing packages, existing under a form or another since a long time, used by many people around the world
- because it provides access to several techniques whose usefulness has traditionally been hindered by the lack of available software: notably, linear response, ultrasoft PP, Car-Parrinello Molecular Dynamics
- because a package that is suitable for easy addition of new developments, extensions, features, is badly needed. Such a goal requires:
 - a modular structure, easy to maintain, easily extensible
 - a collaborative environment: free/open-source license

On the sad status of scientific software in physics

Scientific software used in physics tends to belong to one of the following categories:

- *home-brewed*. Typical for small problems: written from scratch – usually in a amateurish way – and left to “bit-rotting” after publication point is hit.
- *I-have-my-version*. Typical for bulkier problems (e.g. electronic structure): a package or large code is produced by and distributed to a number or people. Everyone makes its own version, with its own set of changes, additions, subtractions, extensions, incompatibilities.
- *don't-touch-me*. Commercial applications, or applications coming with some form of restrictive license. Often no access to sources, no possibility to do any development, and even if you can, no possibility to re-distribute it. Rather common in chemistry and biology, increasingly common in physics.

The alternative: *free/open-source software*

Free software and GNU license

Distributed under many different (and sometimes conflicting) licenses. Most common license:

- GNU (Gnu's Not Unix) General Public License (GPL), inspired by Richard Stallman.

Basically:

- The source code is available.
- You can do whatever you want with the sources, but if you distribute any derived work, you have to distribute under the GPL the sources of the derived work.

Advantages:

- Everybody – including commercial entities – can contribute.
- Nobody can “steal” the code and give nothing back to the community.

The most successful example is probably the Linux Kernel.

Present status

The Democritos PW-PP-DFT package is the result of the merge of three main packages:

- PWscf (Trieste, Lausanne): self-consistent electronic structure, structural relaxation, dynamics, linear-response, postprocessing
- CP (Lausanne, Princeton): variable-cell Car-Parrinello molecular dynamics
- FPMD (Bologna, Trieste): variable-cell Car-Parrinello molecular dynamics

The three packages share common installation method, input format, PP format, data output format, parts of the basic code.

A GUI (Graphical User Interface) for the production of input files is available.

The package is maintained as a single CVS (Concurrent Version System) tree.

Available to everyone anytime via anonymous (read-only) access.

Web site: <http://www.democritos.it/scientific.php>.

Technical characteristics (capabilities)

The Democritos package can be used for

- both Γ -point and \mathbf{k} -point calculation (except CP)
- both insulators and metals, with various flavors of broadening, or tetrahedra (PWscf)
- any crystal structure or supercell form
- both norm-conserving PP's in separable form and ultrasoft Vanderbilt PP's (the latter not yet in FPMD)
- many exchange-correlation functionals (both LDA and gradient-corrected: PW91, PBE, B88-P86, BLYP,...)
- spin-polarized, magnetic systems (also noncolinear for PWscf)

on a large number of machines.

Technical characteristics (algorithms)

- use of iterative techniques: the Hamiltonian is stored as operator, not as matrix. All standard PW technicalities: FFT, dual-space, etc., are used. Iterative diagonalization used whenever it is useful.
- fast “double-grid” implementation for ultrasoft PP’s: the cutoff for the augmentation part can be larger (the corresponding FFT grid denser in real space) than the cutoff for the smooth part of the charge density. CP only: very fast “box grid” implementation.
- Parallelization is performed on both PW’s and FFT grids, using a parallel 3D FFT algorithm having good scaling with the number of processors (memory use also scales)
- Parallelization on \mathbf{k} -points is also available (for PWscf) by dividing the processors into “pools” and dividing \mathbf{k} -points across pools of processors.

Technical characteristics (coding)

- written mostly in Fortran-90, with various degrees of sophistication (i.e. use of advanced f90 features) – no dirty tricks, no parts of code to be ashamed of!
- use of standard library routines (lapack, blas, fftw) to achieve portability – Machine-optimized libraries can (should!) be used if available
- C-style preprocessing options for machine dependencies (e.g. to select machine-optimized libraries) allow to keep a single source tree for all machines
- parallelization via MPI calls, hidden into calls to very few routines – (almost) unified serial and parallel versions. Unless something special is desired, there is no need to know the internals (internals?) of parallelization in order to write parallel code.

Easy (or not-so-difficult) installation via the GNU utility `configure`

Pseudopotentials

PP's must be given in one of the following formats:

- UPF (Unified Pseudopotential Format):
 - formatted (small amount of data)
 - human-readable (may contain info needed to reproduce the PP)
 - extensible: based on “fields” introduced by a XML-like syntax:
 $\langle PP_field \rangle \dots \langle /PP_field \rangle$
 - documentation (sort of) available
 - converters from several some pre-existing formats are available
- Old formats (for compatibility):
 - “Old” norm-conserving PP format, one projector per angular momentum
 - “New” format for both ultrasoft and norm-conserving PP's (more than one projector per angular momentum allowed)
 - David Vanderbilt's format for ultrasoft PP's

How it looks like

<PP_INFO>

Generated using unknown code

Author: Von Barth-Car (<1984)

Info: automatically converted from PWSCF format

0 Generated with a Non-Relativistic Calculation

0.000000000000E+00 Local Potential cutoff radius

nl	pn	l	occ	Rcut	Rcut US
3S	0	0	2.00	0.000000000000	0.000000000000
3P	0	1	2.00	0.000000000000	0.000000000000

</PP_INFO>

<PP_HEADER>

0 Version Number

Si Element

NC Norm - Conserving pseudopotential

F Nonlinear Core Correction

SLA PZ NOGX NOGC PZ Exchange-Correlation functional

4.000000000000 Z valence

0.000000000000 Total energy

0.0000000 0.0000000 Suggested cutoff for wfc and rho

1 Max angular momentum component

431 Number of points in mesh

2 2 Number of Wavefunctions, Number of

Wavefunctions nl l occ

3S 0 2.00

3P 1 2.00

</PP_HEADER>

<PP_MESH>

```
<PP_R>
1.30825992062E-03  1.34137867819E-03  1.37533584110E-03
</PP_R>
<PP_RAB>
1.30825992062E-03  1.34137867819E-03  1.37533584110E-03
</PP_RAB>
</PP_MESH>

<PP_NONLOCAL>
  <PP_BETA>
    1      0          Beta      L
    359
    5.62466109801E-03  5.76705055555E-03  5.91304456932E-03
  </PP_BETA>
</PP_NONLOCAL>
```

CP code

Developed by Alfredo Pasquarello (IRRMA, Lausanne), Kari Laasonen (Oulu), Andrea Trave (UCBerkeley), Roberto Car (Princeton), PG, Nicola Marzari (MIT) and others. Car-Parrinello variable-cell molecular dynamics with Ultrasoft PP's. Main features:

- Verlet dynamics with mass preconditioning
- Temperature control: Nosé thermostat, velocity rescaling
- Metallic systems: Nosé thermostats for both electrons and ions
- electronic and ionic minimization via damped dynamics
- Modified kinetic functional for constant-pressure calculations
- “grid box” for fast treatment of augmentation terms in Ultrasoft PP's

Limitations:

- no k -points
- no fancier minimization schemes
- no constraints

Many developments in the pipeline.

FPMD code

Developed by Carlo Cavazzoni (CINECA, Bologna), Sandro Scandolo (ICTP, Trieste), Guido Chiarotti (SISSA, Trieste), Paolo Focher, Gerardo Ballabio and others. Car-Parrinello variable-cell molecular dynamics. Main features:

- Verlet dynamics with mass preconditioning
- Temperature control: Nosé thermostat, velocity rescaling
- Metallic systems: Nosé thermostats for both electrons and ions
- Modified kinetic functional for constant-pressure calculations
- Various electronic and ionic minimization schemes
- Nudged Elastic Band (NEB) scheme for transition paths and energy barriers
- Constrained dynamics

Limitations:

- no Ultrasoft PP's

PWSCF code

Developed by Stefano Baroni, Stefano de Gironcoli, Andrea Dal Corso (SISSA, Trieste), PG, and others. Main features:

- Self-consistent ground-state energy and Kohn-Sham orbitals
- Structural optimization
- Molecular dynamics on the ground-state Born-Oppenheimer surface
- Variable-cell molecular dynamics
- Phonon frequencies and eigenvectors at a generic wave vector
- Effective charges and dielectric tensors
- Electron-phonon interaction coefficients for metals
- Interatomic force constants in real space
- Third-order anharmonic phonon lifetimes
- Macroscopic polarization via Berry Phase
- Calculation of transition paths and energy barriers using NEB

Limitations:

- no Car-Parrinello dynamics
- very limited constrained minimization and dynamics