Understanding why convergence can be hard

L. D. Marks and R. Luke, October 2006

In most cases the scf cycle in Wien2k converges quickly and cleanly; but in number of cases it does not. These notes are intended to provide some background to why it does not, which may be useful in terms of understanding what to do to improve it in a particular case.

The first key point is to recognize that Wien2k uses an iterative procedure, rather than a variational approach, what is called a “fixed-point iteration”. One starts cycle “n” with some density $\rho_n(r)$, and at the end of a given cycle (in practice, summed from various contributions in mixer) there is a new density $\rho'_n(r)$. Suppose that $\rho_{true}(r)$ is the true electron density that you are trying to obtain. The process of doing a single scf cycle is called contractive iff

$$\| \rho'_n(r) - \rho_{true}(r) \| < \| \rho_n(r) - \rho_{true}(r) \|$$

If the cycle is contractive, the new density is a better approximation and in most cases convergence will be fast. This is not always the case, particularly in the early stages when one starts with the atomic densities from dstart, but it can also occur at other times.

A second key point is the question of whether the densities form a convex set. (Although in principle one can have any density, in reality one is limited to ones which are consistent with the solutions of the Kohm-Sham equations.) Suppose we have two intermediate densities $\rho_1(r)$ and $\rho_2(r)$, the set of possible charge densities from the Kohm-Sham equations including the true solution is called convex iff all $\rho(r)$ defined by

$$\rho(r) = \lambda\, \rho_1(r) + (1-\lambda)\, \rho_2(r),\ 0 < \lambda < 1$$

are also possible charge density from the Kohm-Sham equations. If this condition is not true, the overall problem is what is called non-convex. In this case one often sees what has been called in the literature “charge sloshing”, and in some cases there may even be multiple local minima. In many cases when the problem is non-convex the single scf cycle is also not contractive. The iterations may bounce from one solution regime to another and either converge very slowly or diverge. A non-convex problem is much harder to solve than a convex one. (For reference, the fact that the densities may obey the variational principle does not, alas, guarantee that the overall problem is convex.)

A simple example of a non-convex case is a system with two possible spin states, 1 and 2. For a semiconductor or an insulator with a defect one can also consider solutions where there are electrons in the conduction band verses holes in the valence band as the two possible solutions.

The current mixer often does a very good job of handling cases when the cycles are not contractive and/or the problem is not convex, but this is not always the case and it may need a bit of user help. Some general suggestions:

1) Think about ways to improve the scf calculation, since this may make it more contractive. For instance, the default energies in case.in1 are set at 0.3 which is good for most bulk materials but often too high for a surface. In this case you can set them to about 0.3 Rydberg below the Fermi energy by editing the file by hand. Sometimes a few more k-points can also help; running with just one k-point can give problems. Other things to think about are using more LM's in case.in2, and adding by hand a high energy state to, for instance, remove f-electrons.

2) Check that the Plane Waves and CLM's are reasonably scaled. Bad relative scaling of the two is often shown by a relatively large angle between the step chosen by the Broyden mixing and the Pratt step (grep :DIRB to see this). If this angle is relatively large, e.g. 90 degrees or larger, reducing the PW scaling in case.inm can help. In general, the larger the calculation, the smaller this scaling should be and values of 0.03 are not unreasonable. Unfortunately sometimes the problem is not related to bad scaling of the plane waves but some other type of ill-conditioning and is not cured this way.

3) Be careful about the use of the –in1new switch early in a calculation before the energy and charge are reasonably well converged; this can increase the amplitude of oscillations. You should also check that the number of APW's/LAPW's is not oscillating and adjust the cutoff for choosing these (the –ql switch) if needed.

4) Reduce the amplitude of the Broyden step parameter in case.inm. Going back to the example of two spin states, by doing this you may force the charge density output from the mixer to stay within the set of densities corresponding to charge state 1 rather than wandering into state 2 and oscillating.

5) Almost as a last resort, run the iterations almost by hand until the densities have moved away from a region where they are non-contractive and/or non-convex into a region where at least locally $\rho_n(r)$ and $\rho'_n(r)$ are better behaved. For instance you could use "run_lapw –i 2 …" to just run a couple of iterations, then use save_lapw to save them and keep repeating this until you've reached a stable regime. This way if things blow up you can go back to a more stable intermediate density, adjust some of the parameters and try again.