Modelling crystals from first principles

Tyler C. Sterling[1, *]

[1]*Department of Physics, University of Colorado at Boulder, Boulder, Colorado 80309, USA*

(Dated: August 13, 2021)

This is a **work-in-progess** document intended as an introduction to density functional theory (DFT) for my research group. The idea is to make this an approachable document designed to help my experimentalist-peers become more familiar with the methods and lingo used in computational condensed matter physics. The *flavor* of this project will be lattice dynamics, which is what we study. For my own benefit, I also plan to include the technical details of building a full-potential linearized augmented-projector-wave code (FLAPW or just LAPW) similar to the ELK FORTAN code. I am working on that in parallel to writing this document, so this text will be closely linked with my code.

CONTENTS

I. INTRODUCTION

This document is a work-in-progess project intended to introduce my research group to density functional theory (DFT). The idea is to make this an approachable document designed to help my experimentalist-peers become more familiar with the methods and lingo used in modern computational physics. The *flavor* of this project will be lattice dynamics, which is what we study in my

---

[*] ty.sterling@colorado.edu

research group. To that end, I plan to cover both finite-difference and density functional perturbation theory (DFPT) methods. For my own benefit, I also plan to include the technical details of building a full-potential linearized augemented-projector-wave code (FLAPW or just LAPW) similar to the ELK FORTAN code. I am working on that in parallel to writing this document, so this text will be closely linked with my code.

My end goal is to have a code that can find a ground-state potential that is then frozen and new model-Hamiltonian style potentials can be added; i.e. instead of fitting a tight-binding model to the DFT bands, simply start from a rigid potential and add new terms. This might work as a simple modification to a non-self-consistent (NSCF) calculation or might be more involved. I don't know yet. Furthermore, it may not work at all and I may write a code that is designed to fit tight-binding models. It may make more sense to fit the potential rather than the bands anyway ... I don't know this yet either. To that end, my code will probably depend heavily on reading the potential and wave functions from other DFT packages (e.g. ELK, ABINIT, VASP, ...) since these are much faster and more feature rich than my code ever will be. I am mainly writing the DFT part to get a better grasp of the methods and implementation because I hope to work on one of the packages one day!

A broad outline (*to be updated as I go!*) is as follows (note, this is NOT a table of contents but merely what I plan to put in this document):

1. **What is density-functional-theory**: The Hohenberg-Kohn theorems. The Hohenberg-Kohn theorems prove that the ground-state density of a many-body system can be mapped to a non-interacting problem with the many-body effects included in a mean-field potential (that is an explicit functional of the ground state density). A second theorem also proves that the ground-state density that minimizes the total energy is the true ground-state density and is the only density that does it. *Refs: "Modern Condensed Matter Physics" by Girvin and Yang, "Computational Physics" by Jos Thijssen, "Electronic Structure" by Richard Martin. Probably also the Hohenberg and Kohn papers.*

2. **DFT in practice**: the Kohn-Sham Hamiltonian. The Kohn-Sham (KS) Hamiltonian, also called the KS equations is a set of independent-particle Schrödinger like equations that can derived using the Hohenberg-Kohn theorems. These equations provide the starting point for building a DFT code. Typically a variational basis set is assumed and the KS equations are solved, varying the expansion coefficients until the energy is minimized. *"Computational Physics" by Jos Thijssen, Paolo Gianozzi's lecture notes. Probably need to look at Gold-*

stein and Afken-Weber-Harris to get the Largrange multiplier part right. I think "Modern Consdensed Matter Physics" has an appendix on functional derivatives.

3. **DFT for atoms**: a fundamental task in all (well, most) DFT codes is solving the radial Schrödinger's equation for a spherically symmetric potential. This is the case for atoms, etc. I think this is really only necesssary in muffin-tin type codes (e.g. LAPW, PAW, and USPP methods) and not in pseudopotential (PP) codes... but I am not sure. I will revisit this claim when I get to it. I also should introduce pseudopotentials and other methods here. *Refs: "Computational Physics" by Jos Thijssen, "Electronic Structure" by Richard Martin.*

4. **A brief review of of electronic structure theory**: DFT methods for crystals, like any electronic structure method for periodic external potentials, make extensive use of Bloch's theorem. Use of Bloch's theorem results in it being necessary only to solve the KS equations in the first Brillouin zone of reciprocal space (as opposed to real space in e.g. quantum chemical or molecular methods). Since crystals are assumed to extend over *all* of space, it intractable to solve the equations in real space. Instead, the whole problem is formulated in reciprocal space and the first Brillouin zone is sampled on a fine grid of $\boldsymbol{k}$ points. The total energy E is calculated by integrating the Brillouin zone over all occupied KS eigenvalues $\epsilon_n(\boldsymbol{k})$ *Refs: "Computational Physics" by Jos Thijssen, Vanderbilt's "Berry Phase in Electronic Structure" chapter on Electronic Structure Theory, David Tong's lecture notes..*

5. **DFT for crystals**: the KS equations in a crystal potential. The natural basis set for a reciprocal space code is plane waves (PW). However, due to the requirement of orthogonality of the KS Bloch states $\psi_n(\boldsymbol{k})$, an impractically high number of PWs is required to achieve reasonable precision (remember, the basis set is variational!). Instead, two methods are usually used:

   (a) In the core-region close to the atomic nuclei, local orbitals are used instead of PWs (which are still used in between nuclei), and the PWs are matched to the local orbitals are the core-region boundaries. Thes is the idea behind the APW and LAPW methods as well the old-school, less sophisticated OPW and muffin-tin methods.

   (b) Instead of modifying the wave-functions, we can modify the *potential*. This is the idea behind the pseudopotential (PP) method. In the PP method, the potential due to the density of the core states, as seen by the valence states, is transformed away so that, outside of some specified cutoff region, the valence wave functions are still the same

as they would be if the potential were not modified. In the core-region, the potential is such that the wave function is as smooth as possible, reducing the number of PWs required to get good convergence.

It is also possible to combine both, as in the PAW or mixed-basis pseudopotential methods (cite Rolf Heid's work). The wave functions are transformed *and* the potential is modified so that as few plane waves as possible are required. Generally, the number of PWs required by the methods applying local orbitals are much lower. This is favorable as the KS equation is still expressed in terms of the PWs: the matrix elements are just calculated using the transformation to local orbitals. With fewer basis functions, diagonalizing the Hamiltonian (an operation that scales $\mathcal{O}(N^3)$ is much easier. However, in the case of all electron (AE) methods like LAPW, there are many states present to treat so the computational effort is still large. However, AE methods are the most accurate methods available for crystals. On the otherhand, frozen-core methods like PAW benefit from both fewer basis functions and fewer states in the secular problem. In contrast, PP methods require more basis functions but the codes are much easier to write resulting in more properties being available from PP methods (though PAW and LAPW codes calculating the same properties usually come around eventually). Additionally, pure PWs are a good starting basis for more advanced many-body perturbation methods so PP methods are sometimes used as starting points for those codes. *Refs: "Computational Physics" by Jos Thijssen, "Electronic Structure" by Martin, "Plane Waves and Pseudopotentials" by Singh and Nordstrom, Paolo Gianozzi's lecture notes.*

6. **Review of lattice dynamical calculations**: Here, I will review how the lattice dynamical equations are derived and solved. Ultimately, the goal of a DFT code applied to the lattice dynamical problem is to calculate the inter-atomic force constants (IFCs) as done in the finite-difference method or to calculate the dynamical matrics (which are the space Fourier transform of the IFCs) as done in the DFPT method. *Refs: "Lattice Dynamics" by Dove and "Solid State Physics" by Ashcroft and Mermin. Girvin and Yang do a good job too.*

7. **Finite-diffence**: It is a looong-way away to write this stuff... lets finish the DFT and lattice dynamical calculation parts first.

8. **DFPT**: It is a looong-way away to write this stuff... lets finish the DFT and lattice dynamical calculation parts first. *Refs: "Phonons and related crystal properties from density-functional*

perturbation theory" by Baroni and Paul Tulip's dissertation "Dielectric and Lattice Dynamical Properties of Molecular Crystals via Density Functional Perturbation Theory: Implementation within a First Principles Code" (this thesis does a good job reviewing Gonze's papers deriving the DFPT equations and discusses the implementation of DFPT in the PWPP code CASTEP. *Also Rolf Heid's notes in Pavarini's Correlated-Electons compilations.)*

9. **Electron-phonon coupling**: Second quantization of phonons etc. It is a looong-way away to write this stuff... lets finish the DFT and lattice dynamical calculation parts first. *Refs: "Electron-phonon interactions from first principles" by Feliciano Giustino. Also Rolf Heid's notes in Pavarini's Correlated-Electons compilations.*

That is enough rambling for now. Let's get to work!

II. SCATTERING FROM ATOMS

We will need to know how to solve the Schrödinger's equation for arbitrary (but well behaved!) spherically symmetrical potenials. The goal later on will be to apply this to solving the potential in *core* regions near nuclei in crystals. However, we aren't there yet and numerically solving the radial Schrödinger's equation is a good warm-up to computational physics. So in this section, we will look at a problem we don't care that much about but that gives us a nice example of the connection between computational physics and experiment.

We will study the elastic scattering of (spinless) quantum particles from spherically symmetric potential. We will, in a sense, assume that we already know the scattering potential $V(\boldsymbol{r}) \equiv V(r)$ so that the problem is reduced simply to solving the radial Schrödinger's equation:

$$H\psi(\boldsymbol{r}) = \left[\frac{\hbar^2}{2m}\nabla^2 + V(r)\right]\psi(\boldsymbol{r}) = E\psi(\boldsymbol{r}) \tag{1}$$

This is useful in numerous circumstances: e.g. to solve the KS equations for core-states in a DFT code, for building pseudopotentials, or for studying a real scattering potential in an experiment. In the later case, the resulting eigenvectors ψ and eigenvalues E are used to model measured spectra and the parameters used to to define the potential $V(r)$, are varied to fit the experiment. We will model the scattering of hydrogen off krypton for an example.

As we will see, solving the spherically symmetric Schrödinger's equation ultimately results in numerically solving a 1-dimensional Schrödinger's equation. But before we dive into the numerical solution, lets review the analytical solution for simple cases.

In spherical coordinates, the Schrödinger's reads:

$$-\frac{\hbar^2}{2m}\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\psi}{\partial r}\right)+\frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\psi}{\partial\theta}\right)+\frac{1}{r^2\sin^2\theta}\left(\frac{\partial^2\psi}{\partial\phi^2}\right)\right]+V(r)\psi=E\psi \qquad (2)$$

We can seperate the solution into a radial part and a spherical part: $\psi(\boldsymbol{r})=R(r)Y(\theta,\phi)$. Now we insert this ansatz into eq. 2 and seperate the radial and angular parts to get:

$$\left\{\frac{1}{R}\frac{d}{dr}\left(r^2\frac{dR}{dr}\right)-\frac{2mr^2}{\hbar^2}\left[V(r)-E\right]\right\}+\frac{1}{Y}\left\{\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial Y}{\partial\theta}\right)+\frac{1}{\sin^2\theta}\frac{\partial^2 Y}{\partial\phi^2}\right\}=0 \qquad (3)$$

Since each part is independent of the other, for the right hand side to be true, they must both be equal to $\pm$ the same constant. Anticipating the solution, lets call the constant $l(l+1)$. Then we need to solve the two inhomogeneous equations seperately.

The radial equation is

$$\frac{1}{R}\frac{d}{dr}\left(r^2\frac{dR}{dr}\right)-\frac{2mr^2}{\hbar^2}\left[V(r)-E\right]=l(l+1) \qquad (4)$$

and the angular equation is

$$\frac{1}{Y}\left\{\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial Y}{\partial\theta}\right)+\frac{1}{\sin^2\theta}\frac{\partial^2 Y}{\partial\phi^2}\right\}=-l(l+1) \qquad (5)$$

A. The Angular Equation

The solution to the angular equation follows the standard prescription in all kinds of books and, since the external potential $V(r)$ doesn't appear in the angular part, isn't particularly useful to repeat here. Instead, we just quoute the answer: $Y(\theta,\phi)\equiv Y_l^m(\theta,\phi)\equiv|l,m\rangle$ are the well-known spherical harmonics (see Sakurai's book). Note that they depend on l which shows up in the seperation constant $l(l+1)$ that we chose before; we chose this since it is the eigenvalue of $\boldsymbol{L}^2|l,m\rangle=l(l+1)\hbar^2|l,m\rangle$ which we would have found solving the angular equation anyway.

There is more than one way to define spherical harmonics, but the convention we will use is that the spherical harmonics are normalized and satisfy $\langle l',m'|l,m\rangle=\delta_{l,l'}\delta_{m,m'}$. (See e.g. *Classical Electrodynamics* by Jackson for lots of details). We will quote other properties of the spherical harmonics as we go. I might come back here and update this section if I have to.

B. The Radial Equation

The radial equation is the interesting bit since, as long as the potential is spherically symmetrical, the Schrödinger's equation is seperable and the angular part is always spherical harmonics.

The potential $V(r)$ affects only radial part of the wave function. Let us change variables (again...) and let $u(r) = rR(r)$. Then the radial equation becomes

$$-\frac{\hbar^2}{2m}\frac{d^2u}{dr^2} + \left[V + \frac{\hbar^2}{2m}\frac{l(l+1)}{r^2}\right]u = Eu \tag{6}$$

As promised, this equation is equivalent to a 1-dimensional Schrödinger's equation with wave function u and effective potential $V_{eff}^l(r) = \left[V(r) + \frac{\hbar^2}{2m}\frac{l(l+1)}{r^2}\right]$. The superscript l, which I will drop from now on unless I need it again for some reason, is to denote the dependence on angular momentum. The $\propto l(l+1)/r^2$ term is called the *centrigual term* since it tends to repel particles outwards. Note that since the spherical harmonics are (ortho)-noramlized, $u(r)$ satisfies $\int_0^\infty |u(r)|^2 dr = 1$.

Without specifying properties of the potential, this is as far as we can go. However, if we make the assumption $r^2V(r) \to 0$ as $r \to 0$, (i.e. the potential diverges more slowly than the centrigual term), then for small r, $\frac{d^2u}{dr^2} \sim \frac{l(l+1)}{r^2}$. The solution is $u \sim Ar^{l+1} + Br^{-l}$. To avoid diverging wave functions (and probabilities!), we have to discard the $\sim r^{-l}$ solution so that $u \propto r^{l+1}$. Recalling that $R = u/r$, we see that $R \propto r^l$ which means that aside from s angular momentum states, larger angular momentum tends to repel the wave function from the origin (or core, scattering center, etc.). On the otherhand, if we look far from the origin (assuming the potential is localized), we find $\frac{d^2u}{dr^2} \sim \kappa^2 u$ where $\kappa^2 = 2mE/\hbar^2$. To conserve probability, the solution is $u \propto \exp(-\kappa r)$.

C. Example of Spherical Potentials

Now let us look at some examples of spherical potentials. We will study no-potential ($V(r) = 0$, i.e. the free particle) and the infinite spherical well ($V(r) = \infty$ for $r \geq a$). Let us look at the free particle case first. Inspired by the free particle in cartesian coordinates, we can define $E = \hbar^2k^2/2m$ and change variables so that $\rho = kr$. The radial equation for u becomes

$$\frac{d^2R}{d\rho^2} + \frac{2}{\rho}\frac{dR}{d\rho} + \left[1 - \frac{l(l+1)}{\rho^2}\right]R = 0 \tag{7}$$

This is the *Bessel equation* and the solutions are available willy-nilly in books or you can solve it yourself like I did in the math-methods class (see e.g. *Advanced mathematical methods for scientists and engineers* by Bender and Orszag for how!). The solutions in this context are called the *spherical Bessel functions* and are

$$j_l(\rho) = (-\rho)^l \left[\frac{1}{\rho}\frac{d}{d\rho}\right]^l \left(\frac{\sin\rho}{\rho}\right) \tag{8}$$

and

$$n_l(\rho) = -(-\rho)^l \left[\frac{1}{\rho}\frac{d}{d\rho}\right]^l \left(\frac{\cos\rho}{\rho}\right) \tag{9}$$

As $\rho \to 0$, $j_l(\rho) \to \rho^l$ and $n_l(\rho) \to \rho^{-l-1}$ so that only the j_l solution is kept if the origin is included in the problem. For the infinite spherical well, we require that $j_l(ka) = 0$, which gives $E_{l=0} = \frac{\hbar^2}{2ma^2} \left[\pi^2, (2\pi)^2, ... \right]$. For higher l, a computer can be used to invert the Bessel functions to fing e.g. $E_{l=1} = \frac{\hbar^2}{2ma^2} \left[(4.49)^2, (7.73)^2, ... \right]$.

D. Numerical Solution to the Radial Schrödinger's Equation

We could go on to solve the hydrogen atom and other wacky potentials by hand like is done in quantum textbooks, or we could move on to something that isn't boring and solve the radial equation on a computer. Let's do that!

There are apparently many different numerical integration methods for solving the radial Schrödinger's equation(let's call it the SE from now). Generally, these methods assume that space can be discretized into a finite grid and, assuming initial conditions at a few points, the solution can be solved by integrating across the domain. The method I keep coming across is *Numerov's method*, which is allegedly pretty accurate while being very simple. Another one I have seen is the *Runge-Kutta method* (well actually it's a class of methods) that is attractive because the discretized grid can be adjusted willy-nilly. This is an example of an 'adaptive grid' that can be adjusted on-the-fly to be fine in regions where the solutions vary rapidly, but coarse in regions where the solutions are smooth.

Numerov's method is applicate to (among others) differential equations of the form

$$\frac{d^2y}{dx^2} = -g(x)y(x) + s(x) \tag{10}$$

The SE is an example of this with $s(x) = 0$ and $g(x) = 2m/\hbar^2 \left[E - V(x) \right]$. For the radial equation in 3D, $V(x) = V_{eff}^l(x)$, but this is valid for any 1D SE with well behaved V. Poisson's equation from electricity and magnetism is another example: $d^2\phi(x)/dx^2 = -4\pi\rho(x)$ with ϕ the electrostatic potential and ρ the charge density. In this case, $s(x) = -4\pi\rho(x)$ and $g(x) = 0$.

Suppose our differential equation is defined over some finite regions $x_{min} \leq x \leq x_{max}$. We further assume that the solution decays to negligibly small at x_{max}. We then divide the interval into N points of equal spacing that are Δx wide (it is possible and even useful to use *unevenly* spaced grids, but we will get to that later). We label each point x_i and the corresponding solution etc. by $y_i = y(x_i)$, $g_i = g(x_i)$, and $s_i = s(x_i)$. Then to discretize the differential equation, we Taylor expand the solution around a particular point x_n:

$$y_{n-1} = y_n - y_n^{(1)}\Delta x + \frac{1}{2}y_n^{(2)}(\Delta x)^2 - \frac{1}{6}y_n^{(3)}(\Delta x)^3 + \frac{1}{24}y_n^{(4)}(\Delta x)^4 - \frac{1}{120}y_n^{(5)}(\Delta x)^5 + \mathcal{O}\left[(\Delta x)^6\right] \tag{11}$$

$$y_{n+1} = y_n + y_n^{(1)}\Delta x - \frac{1}{2}y_n^{(2)}(\Delta x)^2 + \frac{1}{6}y_n^{(3)}(\Delta x)^3 - \frac{1}{24}y_n^{(4)}(\Delta x)^4 + \frac{1}{120}y_n^{(5)}(\Delta x)^5 + \mathcal{O}\left[(\Delta x)^6\right] \quad (12)$$

Adding $y_{n+1} + y_{n-1}$, we get

$$y_{n+1} + y_{n-1} = 2y_n + y^{(2)}(\Delta x)^2 + \frac{1}{12}y_n^{(4)}(\Delta x)^4 + \mathcal{O}\left[(\Delta x)^6\right] \quad (13)$$

But we know that $y_n^{(2)} = -g_n y_n + s_n \equiv z_n$. Note that eq. 13 is applicatble to z_n too if replace y by z. If we plug z_n into 13, we get:

$$y_n^{(4)} == \equiv z_n^{(2)} = \frac{z_{n+1} + z_{n-1} - 2z_n}{(\Delta x)^2} + \mathcal{O}\left[(\Delta x)^2\right] \quad (14)$$

Plugging this back into 13, we get

$$y_{n+1} = 2y_n - y_{n-1} + (-g_n y_n + s_n)(\Delta x)^2 +$$
$$\frac{1}{12}(-g_{n+1}y_{n+1} + s_{n+1} - g_{n-1}y_{n-1} + s_{n-1} + 2g_n y_n - 2s_n)(\Delta x)^2 + \mathcal{O}[(\Delta x)^6] \quad (15)$$

and after some manipulation we arrive at *Numerov's formula*:

$$y_{n+1}\left[1 + g_{n+1}\frac{(\Delta x)^2}{12}\right] = 2y_n\left[1 - 5g_n\frac{(\Delta x)^2}{12}\right] - y_{n-1}\left[1 + g_{n-1}\frac{(\Delta x)^2}{12}\right] +$$
$$(s_{n+1} + 10s_n + s_{n-1})\frac{(\Delta x)^2}{12} + \mathcal{O}[(\Delta x)^6] \quad (16)$$

Apparently the error in our integration is order $\mathcal{O}[(\Delta x)^6]$ in the step size. Staring at this equation for little while, we see that if we know y_n and $y_n + 1$, we can iteratively determine the solution at other y_i. Importantly, we can integrate in both the positive and negative x directions. Recall that in the SE, the s_n terms are missing and Numerov's formula is somewhat simpler.

E.   Further reading

- *Numerical Methods in Quantum Mechanics (2019/2020)* by Paolo Gianozzi

- *Introduction to Quantum Mechanics* by David Griffiths

- *Modern Quantum Mechanics* by Jun Sakurai

- *Computational Physics* by Jos Thijssen