Phase diagram of graphene: a mean field analysis

Author: Marc Miranda Riaza
Facultat de Física, Universitat de Barcelona, Diagonal 645, 08028 Barcelona, Spain.

Advisor: Bruno Juliá-Díaz and Alexandre Dauphin

Abstract: We study the electronic structure of graphene and the corresponding phases. We first examine the non-interacting part of the Hamiltonian and show the appearance of Dirac cones. Later we study the full Hamiltonian through Mean-Field theory and show a process of symmetry breaking that leads to the formation of a Charge Density Wave (CDW) phase. We show how Automatic Differentiation, a tool from Machine Learning, can be used as an alternative to the Self-Consistent Field method.

I. INTRODUCTION

Carbon atoms are one the building blocks of numerous materials and molecules. There exist a great amount of allotropes of carbon, such as diamond and graphite, many with important applications in industry. For instance, graphite is used to make pencils and diamonds are used in cutting instruments due to its strength. In graphite, carbon atoms form a 3D crystal. It turns out graphite is composed of stacked 2D carbon crystalline layers that we call graphene. At first, graphene was not believed to exist freely [1]. However, in 2004 Andre Geim and Konstantin Novoselov were able to isolate graphene and in 2010 they were awarded the Nobel Prize in Physics "for groundbreaking experiments regarding the two-dimensional material graphene". Since then, graphene has been the focus of intense research due to its remarkable properties, such as its high electrical and thermal conductivity, as well as its strength [2].

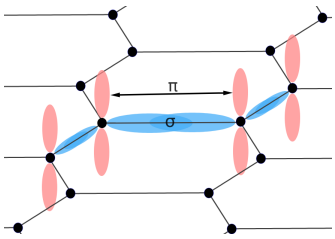


FIG. 1: σ and π bonds in graphene. σ bonds are responsible for the honeycomb lattice. π orbitals are perpendicular to the lattice. For clarity purposes only some bonds have been drawn.

The exceptional properties of graphene come from the outer shell electrons of the carbon atoms, which effectively realize a honeycomb lattice. Therefore, in this work we study the electronic structure of graphene. Carbon atoms have 6 electrons and its fundamental electronic configuration is $1s^2 2s^2 2p^2$. The two innermost $1s$ electrons do not play a significant role in chemical processes. As is depicted in Figure (1), in graphene three of the remaining four electrons bond in a process called sp^2 hybridization with neighbouring carbon atoms. The resulting covalent bonds are called σ-bonds and they gives rise to the honeycomb structure characteristic to graphene. The last electron is in a π-orbital and these bond to form two energy bands, the π-band and the π^*-band. In this work we study electrons in the π bands.

Primarily we are interested in finding the ground state of graphene. We first study non-interacting graphene, which already shows interesting phenomena such as Dirac cones. These are points where the π and π^* bands touch, and they exhibit a linear dispersion relation. The inclusion of Coulomb interactions leads to a rich phase diagram. We use a Mean-Field analysis as a first approximation to the richness of the problem, together with a Self-Consistent Field (SCF) method, an iterative approach to finding the ground state of a Hamiltonian. With this method, we study the phases of graphene. In particular, we show the formation of a Charge-Density-Wave phase.

As an alternative to the SCF method, gradient descent optimization can be used provided that one can efficiently compute gradients. We employ Automatic Differentiation (AD), a core tool from Machine Learning that can yield gradients of complex functions with machine-error precision by applying the chain rule. AD has already demonstrated its usefulness in fields as diverse as Quantum Chemistry [3] and Astronomy [4]. As an example, Tamayo-Mendoza et al. [3] are able to compute the gradient across a SCF calculation with respect to some initial parameters —such as nuclear coordinates and contraction coefficients of atomic orbitals— in a fully-variational Hartree-Fock computation.

This thesis is organized as follows. In section II we discuss the model that is the subject of this work. In section III we examine the non-interacting part of the Hamiltonian and we discuss the non-interacting phase diagram. In section IV we consider the full Hamiltonian and show the interacting phase diagram. In section V we introduce Automatic Differentiation and we discuss how we use it to accelerate the computations. Finally, in section VI we present some conclusions and we discuss possible future work.

II. THE MODEL

We consider that the π electrons can jump between neighbouring sites and interact through the Coulomb force. For simplicity we limit the interaction to nearest-neighbours. In terms of the creation and annihilation operators c_i, $c_i^\dagger$, the Hamiltonian under consideration can be written as

$$H = -t\sum_{\langle i,j\rangle}(c_i^\dagger c_j + H.c.) + V\sum_{\langle i,j\rangle} n_i n_j, \tag{1}$$

where $\langle i,j\rangle$ denotes summation over neighbouring sites and $n_i = c_i^\dagger c_i$ is the number operator. The first term in the Hamiltonian is the so-called hopping term. It is a kinetic term that accounts for the possible movement of electrons between neighbouring sites. The strength of such a process is given by the hopping parameter t. Using a tight-binding approximation, t is found to be around $2.74eV$ [5]. The second term is a nearest-neighbour density-density interaction between electrons located in neighbouring sites. In this work we only consider repulsive interactions, i.e. $V > 0$.

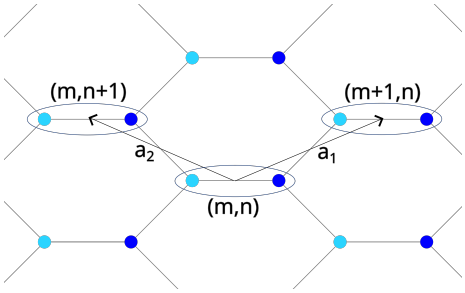


FIG. 2: Honeycomb lattice realized as a hexagonal lattice with a diatomic basis and basis vectors $\{\mathbf{a_1}, \mathbf{a_2}\}$.

Figure (2) shows the geometrical disposition of graphene. The honeycomb lattice can be thought to consist of two hexagonal Bravais sublattices, which we denote ϕ and ψ. Equivalently, the honeycomb lattice is a Bravais lattice with a diatomic basis. Let N_Λ be the number of cells of the Bravais lattice. We have $2N_\Lambda$ electrons (in π bonds) and each energy band has a total of $2N_\Lambda$ states. Therefore, we encounter ourselves in a half-filling situation and in the fundamental state all the electrons will be found in the lower-energy π-band. Using these sublattices the Hamiltonian in (1) —with the addition of a chemical potential difference between both sublattices— can be written as

$$\begin{aligned} H = &-t\sum_{m,n}(c^\dagger_{\psi mn}c_{\phi mn} + c^\dagger_{\psi mn+1}c_{\phi mn} + c^\dagger_{\psi m-1n}c_{\phi mn} + H.c.) \\ &+ V\sum_{m,n}(n_{\psi mn}n_{\phi mn} + n_{\psi mn+1}n_{\phi mn} + n_{\psi m-1n}n_{\phi mn}) \\ &+ \beta\sum_{m,n}(n_{\phi mn} - n_{\psi mn}). \end{aligned} \tag{2}$$

Finally, we introduce the fermionic Fourier operators,

$$c_{\alpha mn} = \frac{1}{\sqrt{N_\Lambda}}\sum_{\mathbf{k}\in 1BZ} e^{i\mathbf{k}\cdot\vec{r}_{mn}}c_\alpha(\mathbf{k}), \tag{3}$$

where $\alpha \in \{\phi,\psi\}$ denotes a given sublattice and $1ZB$ stands for the first Brillouin zone. In the following sections we will use these operators to write the Hamiltonian in wave-vector space, which will turn out to simplify the study of the bulk properties.

III. PROPERTIES OF THE NON-INTERACTING HAMILTONIAN

As a starting point, we first examine the non-interacting part of the Hamiltonian (2). Using the previously introduced Fourier operators, the Hamiltonian can be written as

$$H_t + H_\beta = \sum_{\mathbf{k}} \Psi^\dagger(\mathbf{k}) \begin{pmatrix} \beta & -tA^*_{\mathbf{k}} \\ -tA_{\mathbf{k}} & -\beta \end{pmatrix} \Psi(\mathbf{k}) \tag{4}$$

where $\Psi^\dagger(\mathbf{k}) = (c^\dagger_\phi(\mathbf{k}), c^\dagger_\psi(\mathbf{k}))$ and $A_{\mathbf{k}} = 1 + \exp(i\mathbf{k}\cdot\mathbf{a}_1) + \exp(-i\mathbf{k}\cdot\mathbf{a}a_2)$. The off-diagonal terms correspond to the hopping process between sites ϕ and ψ, while the diagonal ones represent the staggering potential. The use of the Fourier transform allows us to turn to quasi-momentum space, yielding a 2×2 matrix. The Hamiltonian is readily diagonalizable and yields the following band structure:

$$\begin{aligned} \epsilon_\pm(\mathbf{k}) = \pm\sqrt{\beta^2 + t^2|A_{\mathbf{k}}|^2} = \pm[\beta^2 + t^2(3 + 2\cos(\mathbf{k}\cdot\mathbf{a}_1) \\ + 2\cos(\mathbf{k}\cdot\mathbf{a}_2) + 2\cos(\mathbf{k}\cdot(\mathbf{a}_1+\mathbf{a}_2)))]^{1/2} \end{aligned} \tag{5}$$

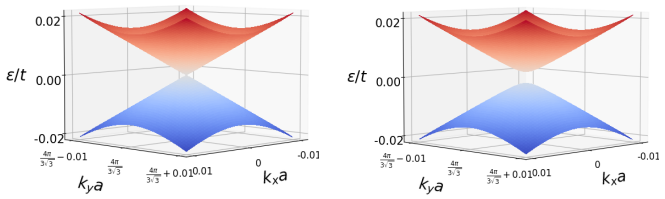
When no staggering is applied ($\beta = 0$) the two bands touch at the so-called Dirac points. At half-filling the lower band is completely filled. Therefore, the Fermi energy of the system is located exactly at band-touching point, this is, at the energy of the Dirac points, which in our case is $\epsilon = 0$. Graphene is here in a semimetallic phase. When a non-zero staggering potential is introduced $\beta \neq 0$, a non-zero band gap appears and graphene behaves as an insulator.

The Dirac points exhibit interesting phenomena, which justifies some further consideration. Their position in wave-vector space is given by those wave-vectors $\mathbf{k}$ satisfying

$$\cos(\frac{3a}{2}k_x) = \pm 1 \quad \text{and} \quad \cos(\frac{\sqrt{3}a}{2}k_y) = \mp\frac{1}{2}. \tag{6}$$

The first Brillouin zone contains 2 non-equivalent Dirac points located at its vertices. One such Dirac point is $\mathbf{k}_D = (0, \frac{4\pi}{3\sqrt{3}}\frac{1}{a})$ and has energy $\epsilon_-(\mathbf{k}_D) = \epsilon_+(\mathbf{k}_D) = 0$. Figure (3) shows the band structure given in (5) in the neighbourhood of the Dirac point $\mathbf{k}_D$. For $\beta = 0$, the

dispersion relation $\epsilon_{\pm}(\mathbf{k})$ is linear, as depicted in Figure (3a). Interestingly, such a linear relationship between energy and momentum is charactheristic of massless ultrarelativistic particles. Thus, electrons around a Dirac cone, named Dirac fermions, behave as if they were ultrarelativistic particles with a different speed than light but with the same behaviour. One has to bear in mind that electrons in bands are no longer free particles, but quasiparticles. If we define a Fermi velocity to be $v_D = \frac{3at}{2\hbar}$, then the energy of the electron would take the usual form for an ultrarelativistic particle $\epsilon(\mathbf{k}) = \pm\hbar v_D|k|$. The Fermi velocity has been found to be approximately $v_F = 1 \cdot 10^6 m/s$, that is, around 0.3% the speed of light. This behaviour suggests that Dirac cones can be used to experimentally study Special Relativity. Indeed, Dirac cones are one example of how condensed matter physics can be used to experimentally test our theories of Special Relativity. Reference [6] shows an example where electrons in graphene exhibit Klein tunneling. Note that the Dirac cone is broken when a non-zero staggering potential is introduced. The quasiparticle mass is related to the curvature of the energy band. Therefore, $\beta \neq 0$ yields massive quasiparticles. As the staggering potential is a chemical potential difference between both sublattices, $\beta > 0$ leads to an increased density in the ψ sublattice and when $\beta \to \infty$, such density tends to 1. When $\beta < 0$ the situation is reversed and the ϕ sublattice is more populated.



(a) (b)

FIG. 3: In (a) Dirac cone located at $\vec{k_D} = (0, \frac{4\pi}{3\sqrt{3}}\frac{1}{a})$, with $\beta = 0$. In (b) the Dirac cone has disappeared with the introduction of a staggering potential $\beta = 0.002t$ that lead to the formation of a gap between both bands.

IV. MEAN-FIELD ANALYSIS OF THE FULL HAMILTONIAN

We now consider the full Hamiltonian. Notice that the hopping term is quadratic on the creation and annihilation operators $c_{\alpha mn}$ and $c^{\dagger}_{\alpha mn}$ —this is, it is a one-body operator—. This allows us to write the Hamiltonian with only a 2×2 matrix. However, the interaction term is quartic —it is a two-body operator—. We can approximate the two-body operator as a one-body operator. In this scenario, each electron interacts with an "effective field" —or mean field— created by the distribution of all the electrons, not directly with the other individual electrons. Therefore, we again have a quadratic Hamiltonian and thus we can proceed in a manner anologous as we did with the hopping term. The Mean-Field method is based on the following Hartree-Fock decoupling approximation:

$$n_i n_j = c_i^{\dagger}c_i c_j^{\dagger}c_j \approx \langle c_i^{\dagger}c_i\rangle c_j^{\dagger}c_j + \langle c_j^{\dagger}c_j\rangle c_i^{\dagger}c_i - \langle c_i^{\dagger}c_i\rangle\langle c_j^{\dagger}c_j\rangle - (\langle c_i^{\dagger}c_j\rangle c_j^{\dagger}c_i + \langle c_j^{\dagger}c_i\rangle c_i^{\dagger}c_j - \langle c_i^{\dagger}c_j\rangle\langle c_j^{\dagger}c_i\rangle) \quad (7)$$

For simplicity we work with ansätze where the density in every sublattice is translationally invariant $\langle c^{\dagger}_{\alpha mn}c_{\alpha mn}\rangle = \langle c^{\dagger}_{\alpha}c_{\alpha}\rangle$ and the expectation value for the hopping term is also translationally invariant, real and rotationally invariant $\langle c^{\dagger}_{\psi mn}c_{\phi mn}\rangle = \langle c^{\dagger}_{\psi mn+1}c_{\phi mn}\rangle = \langle c^{\dagger}_{\psi m-1n}c_{\phi mn}\rangle = \langle c^{\dagger}_{\psi}c_{\phi}\rangle$.

With these considerations, we can write the full Hamiltonian as

$$H = \sum_{\mathbf{k}\in 1BZ} \Psi^{\dagger}(\mathbf{k}) \begin{pmatrix} 3V\langle c^{\dagger}_{\psi}c_{\psi}\rangle & -\tilde{t}A^{*}_{\mathbf{k}} \\ -\tilde{t}A_{\mathbf{k}} & 3V\langle c^{\dagger}_{\phi}c_{\phi}\rangle \end{pmatrix} \Psi(\mathbf{k}) - 3VN_{\Lambda}(\langle c^{\dagger}_{\phi}c_{\phi}\rangle\langle c^{\dagger}_{\psi}c_{\psi}\rangle - \langle c^{\dagger}_{\psi}c_{\phi}\rangle^2), \quad (8)$$

where $\tilde{t} := t + \langle c^{\dagger}_{\psi}c_{\phi}\rangle V$ is a renormalization of the hopping parameter arising from the interaction term. We define the Charge-Density-Wave (CDW) order-parameter to be

$$\rho := \frac{1}{2}(\langle c^{\dagger}_{\phi}c_{\phi}\rangle - \langle c^{\dagger}_{\psi}c_{\psi}\rangle). \quad (9)$$

And we denote the total density as

$$n := \frac{1}{2}(\langle c^{\dagger}_{\phi}c_{\phi}\rangle + \langle c^{\dagger}_{\psi}c_{\psi}\rangle) = \frac{1}{2}. \quad (10)$$

Finally, we denote the expectation value for the hopping process by $\xi = \langle c^{\dagger}_{\psi}c_{\phi}\rangle$. The matrix part of the Hamiltonian can be easily diagonalized, giving again a two-band structure:

$$\epsilon(\mathbf{k}) = 3Vn \pm \sqrt{9V^2\rho^2 + \tilde{t}^2|A_{\mathbf{k}}|^2}. \quad (11)$$

The rest of the Hamiltonian can also be written in terms of n and ρ. The half-filling ground state corresponds to the lower band being completely occupied. Thus, the energy of the maximum state occupied, the Fermi energy, is in this case $\epsilon_F = 0$ and the total free energy per primitive cell is:

$$\frac{F}{N_{\Lambda}} = 3V(-n^2 + \rho^2 + \xi^2) \quad (12)$$

$$+ \int_{1BZ} \frac{v}{4\pi^2}(3Vn - \sqrt{9V^2\rho^2 + \tilde{t}^2|A_{\mathbf{k}}|^2})d\mathbf{k} \quad (13)$$

where $v = \frac{3\sqrt{3}a^2}{2}$ is the volume of a single direct lattice primitive cell and the discrete sum over the states in the first Brillouin zone has been turned to an integral over the Brillouin zone. In Figure (4) we show the free-energy landscape for different relative strengths of the interactions. We can appreciate that for $V/t = 0.5$ the system exhibits a single minimum located at $\rho = 0$, whereas

for $V/t = 2$ there are two symmetric minima with non-zero CDW parameter. Therefore, at some interaction strength the system loses its inherent translational symmetry and forms a CDW phase.

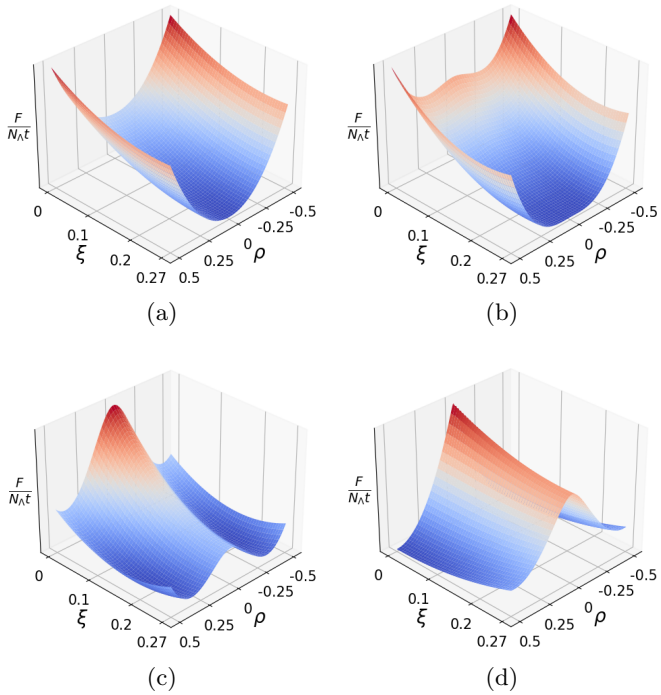


FIG. 4: Free energy landscape for $T = 0$ and 4 different values for $\frac{V}{t}$: 0.5 in (a), 1 in (b), 2 in (c) and 100 in (d).

Now the Hamiltonian itself depends on the wavefunction —through the density and hopping expectation values—. But it is exactly from the Hamiltonian that one calculates the wavefunctions. A possible way out of this difficulty is using an iterative approach. We start with an ansatz wavefunction Φ_0 from which we calculate the necessary expectation values, obtaining the Hamiltonian $H_0(\Phi_0)$. But the original wavefunction is most probably not a solution of this Hamiltonian. Instead, we solve for the ground state Φ_1 of this Hamiltonian H_0 and calculate a new Hamiltonian $H_1(\Phi_1)$ from Φ_1. Repeating this iteration, one hopes that it will eventually converge to a fix point, which would mean that the limiting wavefunction Φ is a solution to the limiting Hamiltonian $H(\Phi)$. We say that Φ is self-consistent, and the iteration method is called the Self-Consistent Field method (SCF).

An example of a convergence sequence is depicted in Figure (5a). Figure (5b) confirms there is a second-order phase transition that leads to the formation of a CDW phase when the interaction is sufficiently strong. If the interaction is not strong enough, the hopping term tends to smooth out an initial asymmetric density distribution between both sublattices. On the contrary, if the interaction is sufficiently strong it tends to magnify small differences between the sublattices. We find that the critical point is between $V/t = 0.92$ and $V/t = 0.93$, which agrees with the reported value of $V/t = 0.924$ given by [7].

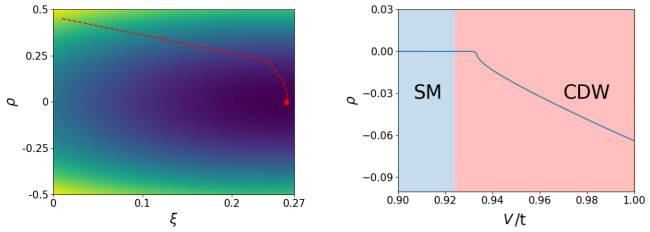


FIG. 5: In (a) convergence path followed by SCF with $\frac{V}{t} = 0.5$ and initial conditions $\xi = 0.01$ and $\rho = 0.45$. In (b) dependence of the CDW order parameter ρ on the strength of the interaction $\frac{V}{t}$. A phase transition between a semimetallic (SM) phase and a CDW phase can be appreciated.The exact transition point depends on the arbitrary choice of when $|\rho|$ is considered large enough to be non-zero. We have selected $V/t = 0.924$ as the critical point.

During our work, the SCF method has not caused any problem. However, the method is known to show convergence problems in some cases [8]. Due to the vital role of the SCF method in many computations [9, 10, 11], improved versions have been proposed. In a damped SCF method [12], the displacement each iteration takes is damped, with the goal of reducing the problem of overshooting. This situation is typical in Gradient-Descent optimization (GD), where a *learning rate* controls the displacement size. This points to the fact that what is important in a SCF method (as well as GD) is the displacement direction —in the (ξ, ρ) plane in our example—. A damping parameter —or learning rate— can then control the displacement size. We denote this damping parameter by α and take it to be between 0 and 1. For $\alpha = 1$ no damping is applied.

V. AUTOMATIC DIFFERENTIATION

Provided we have a way to efficiently calculate the necessary gradients, we could also find the ground state of the system by minimization of the free energy through GD.

Automatic Differentiation (AD) is a tool from Machine Learning that allows one to compute the gradients of complex functions with machine-error precision. It builds a computational graph of all the operations done and uses the knowledge of the derivative of each simple operation together with the chain rule to obtain the final derivative. AD does not come with the high errors associated with numerical differentiation and is more efficient than symbolic differentiation [13]. Furthermore, although we have analytical formulas to compute the gradient in our

case, this is not the general situation. As noted in the introduction, AD has already been used to accelerate computations in Physics and Chemistry [3, 4].

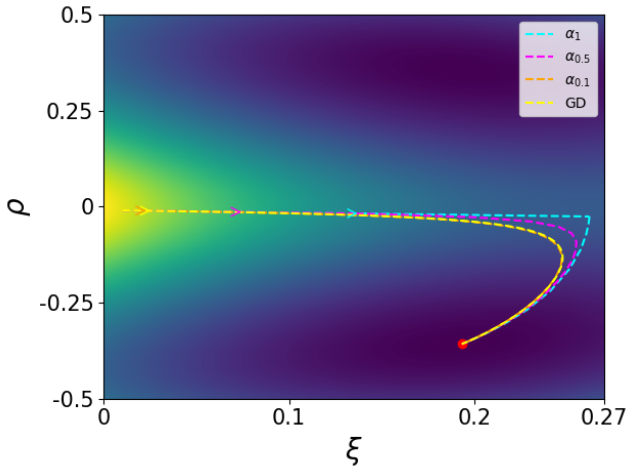


FIG. 6: In yellow a convergence path defined by GD. In other colors, the path followed by SCF for damping values $\alpha \in \{0.1, 0.5, 1\}$.

As depicted in Figure (**??**), the damped SCF path approaches the trail followed by GD as the damping parameter α tends to 0. The same situation was repeated for all calculations. This leads to the conclusion that, at least for the simple problem we are studying, Gradient-Descent and Self-Consistent Field method define the same displacement direction at each point. For a reasonable equal precision, GD with AD was found to be 5 times faster than SCF. As noted previously, the real importance is in the displacement direction and the time it takes to compute. The damping parameter for SCF and the learning rate for GD can be chosen arbitrarily. The time difference between both methods suggests that AD can be used to accelerate computations in places where SCF is used.

VI. CONCLUSIONS AND FUTURE WORK

In this work we studied the electronic structure and phases of graphene. We first examined the non-interacting phase diagram and showed the appearance of Dirac cones. Then, we studied the interacting phase diagram and the formation of a CDW phase through symmetry breaking. Self-Consistent Field is, together with Mean-Field, a common method used in studies like this one.

We showed how Automatic Differentiation can be used to accelerate calculations that involve methods like SCF. In doing so we came to the conclusion that, at least in the conditions of our problem, SCF and Gradient Descent yield the same displacement direction.

There are several ways in which this work could be expanded. Firstly, it would be desirable to theoretically proof that GD and SCF define the same direction. If this is only valid for certain conditions, it would be interesting to determine them. Secondly, although we have shown that AD can be used as a faster alternative to SCF computations, we have done so in the toy example of graphene and using a simple version of SCF. We should consider using state-of-the-art versions of SCF and real-world calculations. Finally, we have only introduced a hopping process and a nearest-neighbour interaction. We should consider adding more phenomena, e.g. next-to-nearest interactions or adding more atoms the the basis.

Acknowledgments

I would like to thank my advisors, Bruno Juliá and Alexandre Dauphin, for all the time they have devoted to this work and for all I have learned from them during this semester.

References

[1] A. H. Castro Neto et al. In: *Rev. Mod. Phys.* 81 (1 Jan. 2009).
[2] Daniel R. Cooper et al. In: *ISRN Condensed Matter Physics* 2012 (Apr. 2012).
[3] Teresa Tamayo-Mendoza et al. In: *ACS Central Science* 4.5 (2018).
[4] Benjamin J. S. Pope et al. In: *The Astrophysical Journal* 907.1 (Jan. 2021).
[5] Rupali Kundu. In: *Modern Physics Letters B* 25 (July 2009).
[6] M. I. Katsnelson, K. S. Novoselov, and A. K. Geim. In: *Nature Physics* 2.9 (Sept. 2006).
[7] A. Dauphin, M. Müller, and M. A. Martin-Delgado. In: *Phys. Rev. A* 86 (5 Nov. 2012).
[8] H. B. Schlegel and J. J. W. McDouall. In: *Computational Advances in Organic Chemistry: Molecular Structure and Reactivity*. Springer Netherlands, 1991.
[9] Su-zhen He et al. In: *Macromolecules* 43.18 (2010).
[10] Jason V. Jorstad, Tian Xie, and Christine M. Morales. In: *International Journal of Quantum Chemistry* 122.10 (2022).
[11] Leon L. Combs and Sten Lunell. In: *International Journal of Quantum Chemistry* 23.2 (1983).
[12] Michael C. Zerner and Michael Hehenberger. In: *Chemical Physics Letters* 62 (1979), pp. 550–554.
[13] Anna Dawid et al. In: *arXiv* (2022).