Dipartimento di Fisica "E. Amaldi"

Università degli Studi Roma Tre

Dottorato in Fisica XIX ciclo

A DFT study of Cr on graphene, with additional material on molecular dynamics.

Michele Monteferrante

Chairperson
of the Ph.D. School
prof. Orlando Ragnisco

Supervisor(s)
prof. Orlando Ragnisco
prof. Antonio Di Carlo

A.A. 2004 - 2005

Summary

Chapter 1: introduction to CNTs and motivations of the work.

Chapter 2: an introduction on molecular dynamics. The Section 2.3 contains a different derivation of what was obtained in [64]. This section contains also a new integration scheme of MD equations.

Chapter 3: a general introduction on Density Functional Theory and pseudopotentials. In Section 3.6 a method to sample semi-classical canonical measures (quantum corrections are expanded in power series of $\hbar$) by fictitious classical MD equations is obtained. This method was applied to a simple quantum system: the harmonic oscillator.

Chapter 4: results of calculations with norm-conserving pseudopotentials. Cr pseudopotential was reconstructed because the library one poorly had reproduced some important physical quantities. All calculations reported in this section were done with the Abinit code.

Chapter 5: results of calculations with ultrasoft pseudopotentials. All calculations reported in this section were done with the CPMD code.

Appendix 1: techniques used to calculate the statistical error in time averaged quantities.

Appendix 2: two definitions of quantities used in the work.

Appendix 3: notations used.

Contents

1 General introduction

1.1 Technological applications of nanotubes

Since their discovery in 1991, carbon nanotubes (CNTs) have received the attention of many scientists. CNTs are obtained rolling up a graphene sheet. The axis along which the sheet is rolled up (see Fig. 1.1) defines the type of nantube.

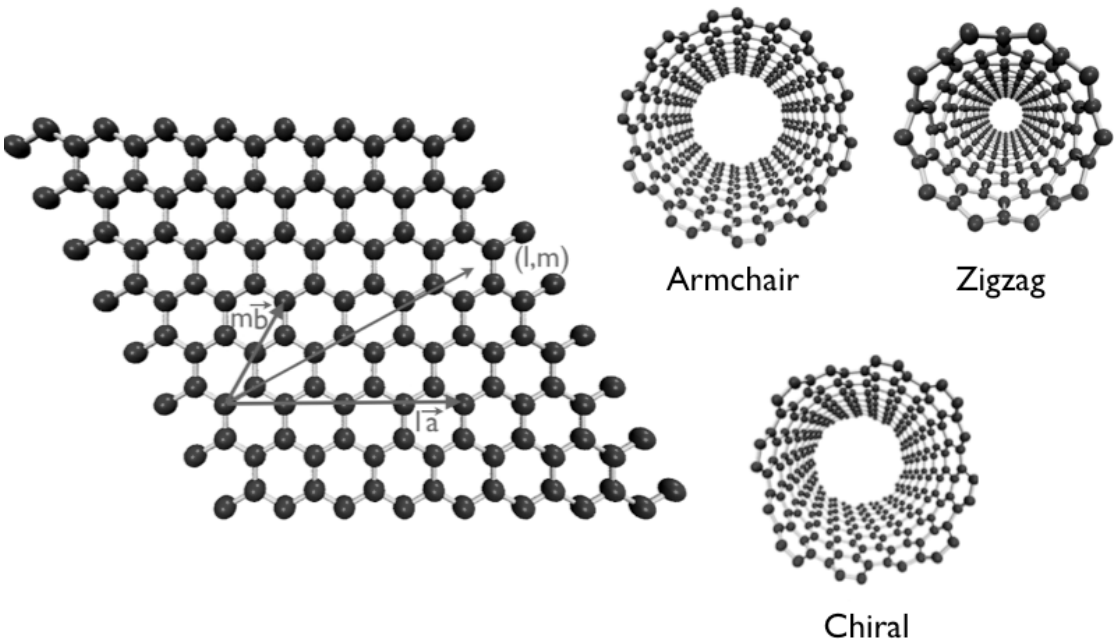


Figure 1.1: Left graphical representation of the (l, m) axis. Right the different types of nanotubes.

(l, l) nanotubes are called armchair nanotubes while $(l, 0)$ are called zigzag nanotube. All other types are called chiral (Fig. 1.1). Multiwall carbon nanotubes (MWCNTs) are composed by a set of concentric nanotubes. The dis-

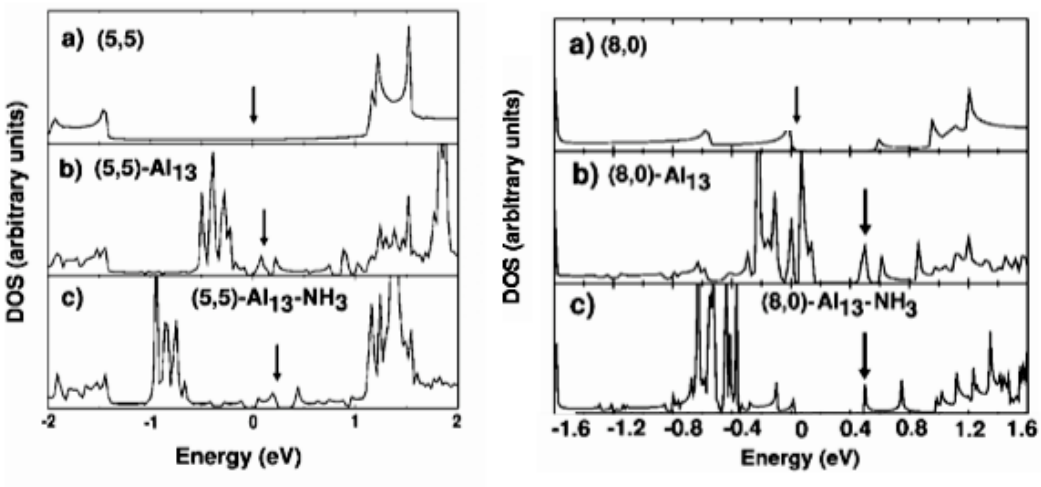


Figure 1.2: Graphics of densities of states from [74]. The arrows indicate the Fermi energies.

tance between adjacent walls is of about 3.41 Å which is the distance between graphite sheets.

Between the various technological applications of CNTs there is the possibility to use them as molecular sensors. In fact even if, due to their similarity to graphite hexagonal network structure, CNTs do not have strong chemical interaction with most molecules, their properties can change with an appropriate chemical environment. For example in [74], through ab-intio MD, was showed how Al_{13} metal cluster can functionalize CNTs. In Fig. 1.2 are reported the densities of states of CNTs functionalized with an Al_{13} cluster. With reference to Fig. 1.2 the capability of the system to detect the presence of an NH_3 molecule absorbed on the Al_{13} it is clear. The effect of Al_{13} is even stronger when an armchair CNT is used because of the stronger interaction between this kind of tube and the metal cluster.

The electronic transport properties of CNTs were also largely studied (see for example [42, 43]). For example, through ab-initio simulations, it was obtained that armchair CNTs are always metal while zigzag tubes, with diameters up to 1.5 nm, are small gap semiconductors. The conductivity property of chiral tubes depend instead on the values of (l, m): a CNT is a metal if $l - m = 3k$ (where k is an integer). Moreover in [42] was also showed that armchair tubes, with a diameter larger than 0.7 nm, keep their metallic character for large bend angles (up to 42°). Conversely, in chiral tubes, the curvature

induces a gap in the band structure which increases as bend angle becomes greater. These studies have suggested to use armchair tubes as flexible conductors and chiral tubes as nanostrain sensors.

The elastic properties of CNTs and MWCNTs are surprising. In general small samples of materials do not keep the same elastic properties of large ones due to the presence of defect and impurities. Conversely nanotubes have a Young modulus and torsional shear modulus very similar to those of graphite and diamond [39, 44, 45, 73]. Various works focused the attention on the possibility of bringing different tubes together Fig. 1.3. In fact since the Van der Waals interaction between adjacent walls is weak the shear modulus of such a nonofiber is very similar to that of a single CNT. The links are produced inducing defects which are produced by electron or ion-beam irradiation (see [33] and references therein). The possibility lo link a large number of CNTs permits the fabrication of nanoropes and nanowires.

Moreover the possibility of reinforcing the MWCNTs torsional shear modulus, producing stronger interactions between adjacent walls, was also studied Fig. 1.3. In [31], classical MD techniques showed that the presence of a low concentration of inter-walls defects can improve the load transfer between shells and increase the interlayer shear strength by several orders of magnitude.

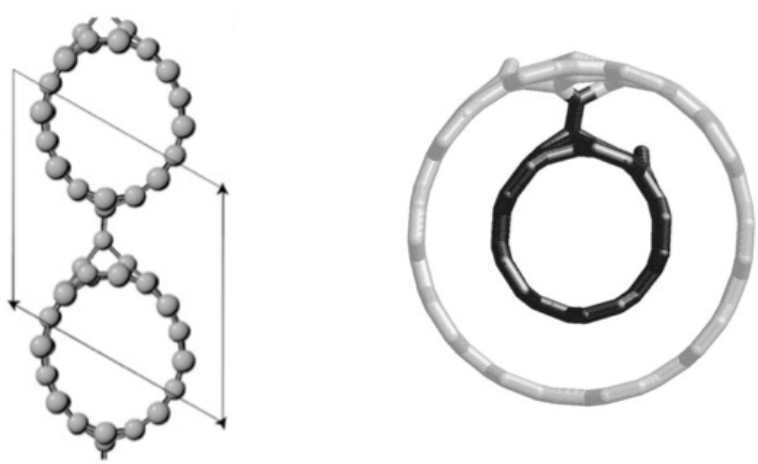


Figure 1.3: Left link between adjacent tubes (figure from [33]). Right link between two CNTs composing a double wall CNT (figure from [31]).

1.2 Motivation of Cr on graphene study.

This work was motivated by results obtained by P.A. Williams et al. [71, 72]. These authors fabricated nanometer-scale mechanical devices ("paddle" oscillators) incorporating multiwalled carbon nanotubes as torsional spring elements (Fig. 1.4a), and developed a method for measuring the torsional stiffness of the nanotubes (Fig. 1.4b). Arc grown MWCNTs were dispersed onto Si wafers having 500 nm of oxide. Large metal pads were patterned by electron-beam lithography over the ends of each MWCNT, and a stripe of metal over its center to form the paddle. The metal was thermally evaporated, 15 nm of Cr followed by 100 nm of Au. The oxide was etched such that the paddles were completely undercut, but the larger pads pinning down the MWCNT ends were not. Measurements revealed a remarkable stiffening behaviour of MWCNTs: after nearly 500 repetitions of twisting cycles of small amplitude (with estimated in-wall strains less than 1%), the stiffness of an individual MWCNT saturated to a value 12 times larger than its initial value.

Young modulus and torsional shear modulus of CNTs and MWCNTs were calculated in [39] using an empirical force-constant model. Young moduli of CNTs were found to be insensitive to the size and helicity. The calculated value was of 970 GPa. The calculated torsional shear moduli of CNTs ranged between 436 GPa (a CNT of 0.34 nm radius) and 478 GPa (a CNT of 13.5 nm radius). In this case only a small dependence on tube radius was found and no dependence on helicity. MWCNTs elastic moduli had the same characteristics:

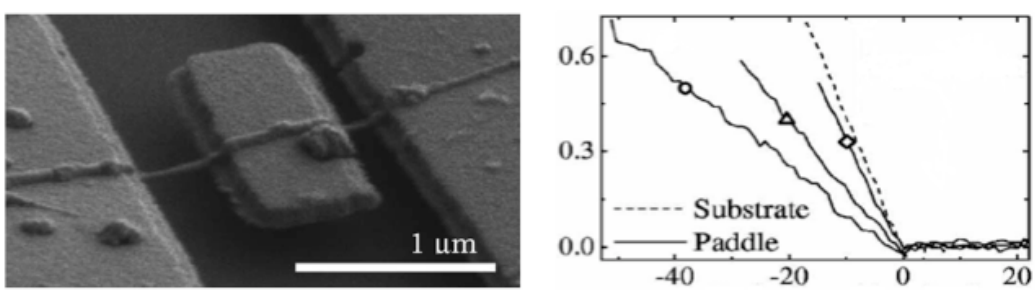


Figure 1.4: a) Paddle oscillator; b) Force-displacement traces on the substrate and on the paddle with three different eccentricities: photodiode signal (nA) vs. piezo-displacement (nm); the slope of the substrate trace yields the apparent overall stiffness of the AFM cantilever, equal to -41 pA/nm (from [48]).

Young modulus either was independent of chirality and equal to 1.1 TPa for all nanotubes with outer shell radius between 1.7 nm and 3.4 nm. Torsional shear moduli were independent of chirality and ranged between 0.455 GPa (a double wall CNT with outer shell radius of 0.68 nm) and 541 GPa (a MWCNT of 10 shell with outer shells radius of 3.39 nm). It is remarkable that elastic properties of CNTs are similar to those in diamond and graphite. Torsional shear moduli of MWCNTs were calculated in [39] assuming a constant value of torsional angle for all shells of MWCNTs, that is as if the shells were strongly interacting between themselves. Conversely the walls interact only through Van der Waals which should not be able to produce a similar inter-shells register.

The outcome of results reported in [71, 72], described at beginning of this section, indicate that an ever stronger inter-walls coupling appears during twisting cycles of the paddle. Results are quantitatively in agreement with this hypothesis as shown with a continuum model [14]. This fact demonstrates that initially Chromium only interacts with outer walls, since, if this was not the case, also in the first twisting cycles all walls should contribute to the torsional shear modulus.

In [60] various types of defect which could increase the inter-shells coupling in MWCNTs were examined: two close vacancies in different graphene sheets, a single vacancy, an interstitial carbon atom (see Fig. 4.2). All these defects have a formation energy greater than 7 eV and then an even greater activation energy. It is not possible that these defects form during the twisting cycles which MWCNTs undergo since their formation energy is too high. Moreover

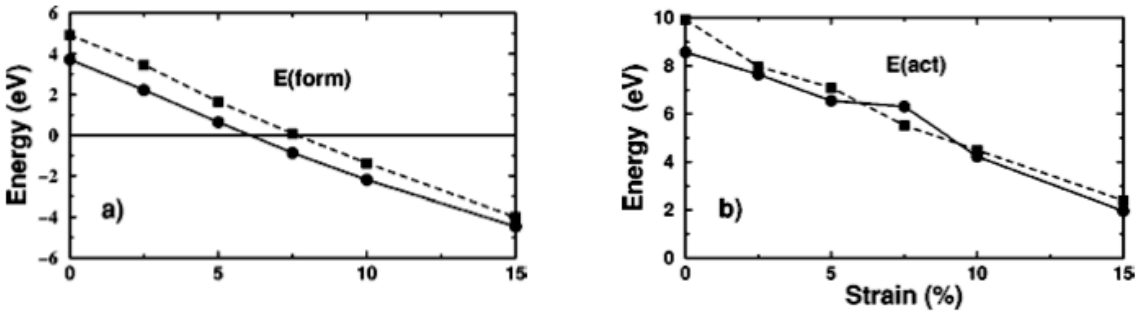


Figure 1.5: Figure from [73]. a) Formation energy of the (5-7-7-5) in a (5,5) SWCNT (circles) and a graphene sheet (squares) as a function of uniaxial. b) Activation energy for the formation of the (5-7-7-5) defect as a function of uniaxial.

Device	Diameter (nm)	F_0 (MHz)	K_θ (10^{-4} Nm)	G_s	G
1	12	1.68	0.37	**570**	90
2	27	2.37	1.6	**460**	35
3	28	2.50	1.9	**740**	54
4	35	3.27	3.6	**560**	33
5	16	2.92	2.4	3420	**435**
6	20	3.79	2.6	2880	**290**
7	18	4.12	7.4	5000	**565**
8	22	2.98	3.9	2190	**200**
9	21	2.04	1.0	1200	120

Table 1.1: Experimental measurements from [49] .

in [73] the stability and activation energies of Stone-Wales defect under tensile strain was calculated Fig. 1.5. SW defect has a formation energy similar to vacancy and of about 7 eV. As shown in Fig. 1.5 even if the SW defect becomes stable for strain greater than 7.5% the activation energies remain high. Moreover in the situation in which we are interested in, the in-wall strains are less than 1%. It doesn't seem possible that these mechanisms produce the described stiffening behaviour of MWCNTs [71, 72] and so Chromium-MWCNT interaction has to play an important role in the observed stiffening.

Other measurements of the resonant frequencies of MWCNTs, fabricated in the same way, indicate that the stiffening does not happen for all MWCNTs. In [49] the resonant frequency F_0, the diameter D and the torsional spring constant K_θ of different MWCNTs were measured. From this data, through the relation $K_\theta = (2\pi F_0)^2 I$ (where I is the inertia moment), the authors obtained an estimate of G_s and G, respectively, the torsional shear modulus for the "outer shell model" (only the outer shell contributes to the shear modulus) and the "solid row model" Tab. 1.1. The error on I was of about 20% which results in an error of about factor 2 for G_s and G. Confronting the two estimated limit values of G_s and G with the theoretical one (about 500 GPa [39]) it was found that one of the limiting cases was too large or too small. For devices 1-4 the "outer shell" model fits while for devices 5-8 fits the "solid rod" one. Device 9 was found in an intermediate situation Tab. 1.1. This suggests that the stiffening of MWCNTs can be produced from a cascade of chemical reactions which can be activated by twisting cycles or not.

This conjecture is also confirmed by the observations of other experimentalists which manipulate MWCNTs in a similar way [19] to fabricate nano-

electromechanical devices where an individual MWCNT, suitably engineered, provides a rotary bearing in which the sliding occurs between different walls. The interwall friction is extremely small and does not increase during operation: beyond doubt, the mechanism producing the stiffening is not in action. It is reasonable then to think that Van der Waals interactions are not the explanation of the stiffening behaviour because in this case the stiffening should be observed. Interestingly, the most effective technique for producing a nanorotor seems to be mounting a metal plate on a MWCNT as in Fig. 1.4a, and then breaking the outer wall off the anchors by a few twisting cycles of large amplitude [19]. This fact prompts the conjecture that the interaction between the CNT and the Chromium is crucial in determining both the fragile behaviour under a few large twists and the ratcheting effect under many small ones. This interaction is far from being understood.

1.3 Aim and techniques

The aim of this work was to characterize the behaviour of Cr evaporated onto CNTs to fabricate devices such as the one described in the previous Section. In particular we tried to understand in what way the presence of Chromium should facilitate the creation of defects able to link the shells of a MWCNT. Moreover one would also like to know in what way the Chromium can penetrate the MWCNT shells. The activation energies of such processes have to be quite small to be activated at room temperature. Finally it should be interesting to understand why these chemical reactions are activated in a cascade way.

We decided to use DFT ab-initio techniques. In these techniques forces between atoms are calculated resolving the electronic problem at fixed ion positions. The discovery of DFT and related Kohn-Sham method has permitted the practical utilization of ab-initio techniques which however remain time consuming. However even to use these techniques many approximations have to be done to reduce the computational cost and the validity of these approximations have to be checked. I refer, in particular to the pseudopotential (Sec. 3.4), number of basis wavefunctions (Sec. 3.4), exchange-correlation functional (Sec. 3.2) and so on. Both norm-conserving and ultrasoft pseudopotentials were used respectively with Abinit and CPMD package. This work was done in collaboration with A. Di Carlo and F. Cleri.

1.4 Additional material

This thesis contains also some additional works regarding molecular-dynamics. In Sec. 2.3 I used different techniques to obtain the result obtained in [64]. In this article differential-geometry techniques were used to show why the Nosé-Hoover thermostat produce the canonical ensemble. I used instead a simple theorem on the invariant measures of first order differential equations.

Sec. 2.6 a new algorithm to integrate MD equations was derived. The aim of this algorithm is to overcome the problem of predictor-corrector algorithms which allow a larger time-step to integrate equations of motion but are not symplectic.

In Sec. 3.6 a way to sample semiclassical (expanded in power series of $\hbar$) canonical partition-function was derived. The application of this techniques is limited to high temperatures. Sections 2.6, 3.6 were developed in collaboration with G. Ciccotti. Possible applications are reported in the Conclusions.

2 Introduction to molecular dynamics

To describe the physical behaviour of large size systems, time averaged values of physical phase-space (x are the phase-space points) observables are essential to characterize the system:

$$\langle A(x) \rangle = \lim_{t \to \infty} \int_0^t A(x(\tau)) d\tau \tag{2.1}$$

A more detailed information on the system, for example velocity and position of particles, are experimentally inaccessible, so it is of no utility to take care of this information in constructing a theory of such systems. In other words, for large size systems, we are interested in statistical properties.

Thermodynamics is the phenomenological theory describing macroscopic properties of large size system and is based on some fundamental principles. Statistical mechanics gives a theoretical foundation to the thermodynamics of reversible processes and, in these situations, represents the link between microscopic and macroscopic quantities. For example in thermodynamics there are some quantities, like the specific heat or compressibility, which can be measured applying a specific transformation to the system. SM gives us the possibility to calculate these quantities, just by knowing the type of interactions between the particles composing the system of interest. In SM time averages of phase space observables are substituted by weighted time-independent averages over phase space accessible to the system (let us call the accessible phase space D):

$$\langle A(x) \rangle = \int_D A(x) d\mu(x) \tag{2.2}$$

These averages are called ensemble averages. The specific measure expression depends on the thermodynamic variables which one decides to fix (N, V, E), (N, V, T), etc ..

Under general hypotheses, one can effectively construct the measures in ensembles like (N, V, T) and (N, P, T), from the knowledge of the measure of (N, V, E) ensemble. The measure of (N, V, E) ensemble can be reasonably constructed using the Liouville's theorem.

Unfortunately, calculations which must be done to have physical information from SM cannot be done by hand and computer simulations become an important tool in studying large size systems. Computational methods which use the measure of interest expression are called Monte Carlo methods Eq. 2.2.

However in a computer simulation we can calculate velocities and positions of all particles and then use them to make time averages. These averages have to be equal to the ensemble average of SM if the ergodic hypothesis, on which SM is based, is satisfied. This averaging technique is called Molecular Dynamics Eq. 2.1. Using MC techniques, if one is interested in dynamic quantities, equations of motion have to be unavoidably solved.

Two classic textbooks on MD and MC methods are [2, 20].

2.1 Liouville's theorem, invariant measures and ergodic theory

Let us suppose to deal with a system characterized by phase-space points x, an evolution law(a flow: T_t) generated by a set of first order differential equations and a probability density distribution function $f(x,t)$. $f(x,t)$ represents the probability to find the system in the state x at time t. If $f(x,0)$ is the probability distribution at time $t = 0$, the probability at time t is obtained from $f(x,0)$ evolving x until t.

The integral of f over phase space points D accesible to the system must be equal to 1 for any t because $f(x,t)$ is a probability density distribution, that is:

$$\int_D f(x,t)dx = 1 \tag{2.3}$$

from now on the domain of integration D will be omitted. it is easy to show that if Eq. 2.3 is satisfied, $f(x,t)$ has to obey the following equation:

$$\frac{\partial}{\partial t} f(x(t),t) + \nabla_x[\dot{x} f(x(t),t)] = 0 \tag{2.4}$$

where $x(t)$ is the evolution of x under the flow. If one is interested in equilibrium distribution functions, f cannot depend explicitly on t, so Eq. 2.4 becomes:

$$\nabla_x[\dot{x} f(x(t),t)] = 0 \tag{2.5}$$

if $f(x,t)$ satisfies Eq. 2.5 then it generates an invariant measure:

$$\mu(T_t X) = \int_{T_t X} f(y)dy = \int_X f(y)dy = \mu(X) \tag{2.6}$$

where $X \subset D$. It can be shown that a necessary and sufficient condition, so that $f(x)dx$ is an invariant measure, is that $f(x)$ satisfies Eq. 2.5 (see [12]). Eq. 2.5 will be used later to see how the Canonical ensemble is generated via MD.

There is a branch of mathematical physics, called ergodic theory, which concerns the study of abstract dynamical systems which possess an invariant measure (see for example [12]).

Suppose to have a set of objects (μ, Ω, T), where μ is an invariant measure, T a discrete-time evolution law and Ω a probability space. Since the evolution

law is of discrete-type, given a point x at time zero, it will be at point $T^k x$ at time k. The measure is invariant in the same sense of Eq. 2.6, that is:

$$\mu(TX) = \mu(X) \tag{2.7}$$

The Birkoff-Khinchin theorem is of central importance in ergodic-theory. Under general hypotheses, this theorem asserts that, if $g \in L_1(\Omega, \mu)$, then:

$$\exists\, \bar{g}(x) \in L_1(\Omega, \mu) : \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n} g(T^k x) = \bar{g}(x) \quad \mu - a.e \tag{2.8}$$

$$\overline{g}(Tx) = \overline{g}(x) \quad \mu - a.e. \tag{2.9}$$

$$\int \overline{g}(x) d\mu(x) = \int g(x) d\mu(x) \tag{2.10}$$

Ergodic dynamical -systems are those systems for which the average over time of an observable $\overline{g}(x)$ does not depend on x, than from Eq. 2.10:

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n} g(T^k x) = \int g(x) d\mu(x) \tag{2.11}$$

that is, time averages are equal to ensemble-averages $\mu - a.e.$

There are many equivalent definitions of ergodic dynamical-systems. For example it can be shown that a dynamical system is ergodic, if and only if, for any $A \in \Omega$ with $\mu(A) > 0$, it results $\mu(\cup_n T^n A) = 1$. This means that starting from any set A with measure different from zero, the trajectory covers the full space Ω. From this point of view it is easy to understand why a dynamical system with a symmetry cannot be ergodic. In fact, wherever the system starts it cannot leave the set of points which respect the symmetry and therefore cannot cover the full space. It can be shown (see for example [18]) that, if a dynamical system has not first integrals ($\mu - a.e.$) of motion independent from μ, it is ergodic.

Under appropriate regularity hypotheses, Poincaré showed (see [18]) that quasi-integrable dynamical systems (an integrable system plus an $O(\epsilon)$ perturbation) cannot have regular first integrals of motion independent of the total energy of the system (H). Following this theorem Fermi tried to show that all quasi-integrable systems with more then two degrees of freedom are ergodic. In fact Fermi showed that, under the same hypotheses of the Poincarè theorem, for such systems the surface $H = cost$ cannot be decomposed into two surfaces

with a smooth boundary. Such a frontier should be a first integral of motion independent on H.

However, as was noted by Urbanski and recognized by Fermi, Fermi's theorem does not exclude the presence of an irregular boundary.

Paradoxically Kolmogorov-Arnold-Moser theorem prooved that, given a sufficiently small ϵ, exists a set of positive measure starting from which the system doesn't cover the full phase space (invariant tori), then it is not ergodic. SM however, also gives good results for systems which are not ergodic (for example a gas of noninteracting particles).

2.2 Hamiltonian dynamical systems and microcanonical ensemble

The statistical ensemble of an isolated N particles system is called "microcanonical ensemble". Isolated classical physical systems are described, in the formalism of analytical mechanics, by Hamiltonian equations. Suppose to have a classical system in three dimensions. Let (q_i, p_i) (with $i = 1, .., 3N$) be the positions and momenta of particles and $H(q_i, p_i)$ the total energy of the system. Equations of motion in the phase-space (q_i, p_i) can be written as:

$$\begin{aligned} \dot{q}_i &= \frac{\partial H}{\partial p_i} = \{q_i, H\} \\ \dot{p}_i &= -\frac{\partial H}{\partial q_i} = \{p_i, H\} \end{aligned} \tag{2.12}$$

where $\{f, H\}$ is the Poisson bracket of f with H:

$$\{f, g\} = \sum_{i=1}^{3N} (\frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i} - \frac{\partial g}{\partial q_i}\frac{\partial f}{\partial p_i}) \tag{2.13}$$

It can be shown that g is a constant of motion if and only if $\{g, H\} = 0$.

Let us return to Eq. 2.6 and let $f(q_i, p_i) \prod dq_i dp_i = f(q_i, p_i) d\Gamma$ be an invariant measure. In the case of Hamiltonian dynamical systems Eq. 2.6 writes:

$$\{f, H\} = 0 \tag{2.14}$$

i.e., f can depend only on the constants of motion, that is if we suppose that H is the only constant of motion, $f = f(H)$.

As we shall see later, in a computer simulation of an N classical particles system, we have to solve Eq. 2.12 as to conserve, at least approximately, constants of motion. In other words, if $(q(t), p(t))$ is the approximate solution of equations of motion and at time t and $H(q(0), p(0)) = E$ then:

$$H(q(t), p(t)) = E + O(\epsilon) \quad for\, all\, t \tag{2.15}$$

There are also discrete symmetries which have to be preserved by the approximate dynamics, like time-reversal. This mean that in a computer simulation, if one decides to solve Eq. 2.12, the invariant measure f can only have the form:

$$f = \begin{cases} constant & \text{in } \Gamma_{E+\epsilon} \quad \Gamma_{E+\epsilon} = \{\Gamma : E < E' < E + \epsilon\} \\ 0 & \text{elsewhere} \end{cases}$$

from wherever value of E we start the simulation.

However, as we have seen in the previous section, a system with symmetries cannot cover the full space Γ. This means that partition function has to be written as:

$$Q(N, V, E) = \int_{E<E'<E+\epsilon} f(q, p) d\Gamma = \int_{E<E'<E+\epsilon} \delta(H(q, p) - E') d\Gamma \tag{2.16}$$

where V is the space-volume in which the system is confined. It can be shown that the measure $\delta(H - E')_{E<E'<E+\epsilon} d\Gamma$ admits a limit for small ϵ:

$$d\mu = \lim_{\epsilon \to 0} \delta(H - E') d\Gamma_{E<E'<E+\epsilon} \propto \frac{d\Sigma}{||\nabla H||} \tag{2.17}$$

If the system is ergodic (if there aren't other symmetries independent of H) time averages can be replaced by ensemble averages. The calculation of partition function is of central importance in statistical mechanics and all thermodynamics quantities can obtained by it:

$$S(N, V, E) = k_B \ln Q(N, V, E)$$

$$1/T = (\tfrac{\partial S}{\partial V})_{E,N}$$

$$P = T(\tfrac{\partial S}{\partial E})_{V,N} \tag{2.18}$$

The equilibrium distribution function in canonical ensemble (see [30]) can be obtained dividing the total system of N particles into two systems: one with N_1 particles and the other with N_2. Suppose that the total system is microcanonical and that $N_1 \ll N_2$. If E is the energy of the whole system, than $E = E_1 + E_2$ with $E_1 \ll E_2$. Since the total system is microcanonical the probability to find the system 1 in a state with energy E_1 must be:

$$f_{N_1}(q_1, p_1, E - E_1) \propto Q_2(E - E_1) = \int_{E_2 \simeq E - E_1} f_N(q_1, p_1, q_2, p_2) d\Gamma_2 \tag{2.19}$$

where f_N is the microcanonical density probability function. Using relations 2.18 and the fact that $E \gg E_1$ we can expand f_{N_1} near E obtaining:

$$f_{N_1} \propto e^{-H_1/kT} \tag{2.20}$$

This is the equilibrium distribution function of the canonical ensemble or (N, V, T) ensemble. In a similar way equilibrium distribution functions in other ensembles can be obtained [30].

The way in which the canonical ensemble has been obtained from microcanonical ensemble is of no practical utility in a computer simulation. In fact, we should simulate $N_2 + N_1$ particles to extract a statistics of only N_1 particles with $N_2 \gg N_1$. In general in a computer simulation one has to modify the Hamiltonian equations of motion adding thermostats such that equilibrium distribution function of the system only, (after integrating out the thermostat degrees of freedom) is of the desired type.

2.3 Non-Hamiltonian dynamical systems: canonical and constrained ensemble

In [28, 46] non-Hamiltonian equations of motion which couple the system to a thermostat (Nosé-Hoover thermostat) to generate a canonical ensemble were introduced. A similar procedure can also be used to generate (N, P, T) ensemble [29].

In this section we show explicitly how Nosé-Hoover thermostat generates a canonical partition function [64]. A different derivation can be found in [64] and a more geometrical treatment of the problem in [16]. Moreover we derive partition function of a constrained system [64]. Partition function of constrained systems will be used in Sec. 2.7. Nosé-Hoover equations write:

$$\begin{aligned} \dot{r}_i &= \frac{p_i}{m_i} \\ \dot{p}_i &= F_i - \frac{p_\eta}{Q} p_i \\ \dot{\eta} &= \frac{p_\eta}{Q} \\ \dot{p}_\eta &= \sum_{i=1}^{dN} \frac{p_i^2}{m_i} - L k_B T \end{aligned} \tag{2.21}$$

where L is a parameter (to be determined), d is the spatial dimension, F_i are the forces and Q determines the thermostat's time-scale. (η, p_η) are the thermostat's degrees of freedom and $i = 1, ..., dN$. it is easy to verify that the function:

$$\begin{aligned} H' &= \sum_{i=1}^{dN} \frac{p_i^2}{2m_i} + V(r_i) + \frac{p_\eta^2}{2Q} + L k_B T \eta \\ &= H(p_i, r_i) + \frac{p_\eta^2}{2Q} + L k_B T \eta \end{aligned} \tag{2.22}$$

is a constant of motion. To obtain the invariant measure we have to solve Eq. 2.5 which writes:

$$\sum_i \nabla_{r_i}[(p_i f(\eta, r_i, p_\eta, p_i)/m_i)] + \sum_i [\nabla_{p_i}(F_i - p_\eta p_i/Q) f(\eta, r_i, p_\eta, p_i)/m_i)] +$$

$$
\begin{aligned}
&+\partial_\eta[p_\eta f(\eta, r_i, p_\eta, p_i)/Q] + \partial_{p_\eta}[(\sum_i \frac{p_i^2}{m_i} - Lk_BT) f(\eta, r_i, p_\eta, p_i)] = \\
&= \sum_i [p_i/m_i \nabla_{r_i} + (F_i - p_\eta p_i/Q)\nabla_{p_i}] f(\eta, r_i, p_\eta, p_i) + \\
&+[p_\eta \nabla_\eta + (\sum_i \frac{p_i^2}{m_i} - Lk_BT)\nabla_{p_\eta}] f(\eta, r_i, p_\eta, p_i) = N d\, p_\eta f(\eta, r_i, p_\eta, p_i)/Q = \\
&= \dot{x}\nabla_x f = dN p_\eta f/Q \quad where \;\; x = (r_i, p_i, \eta, p_\eta)
\end{aligned}
\tag{2.23}
$$

In analogy to what has been said in section 2.2 for Hamiltonian dynamical systems, in a MD simulation of Eq. 2.21, only phase-space points which belong to $H' = constant$ have to be sampled. It follows that the invariant density measure has to be written as:

$$
f = \delta(H' - C) f_2(r_i, p_i, \eta.p_\eta) = f_1(H') f_2(r_i, p_i, \eta.p_\eta) \tag{2.24}
$$

Inserting f in Eq. 2.23 we find:

$$
f_2 \dot{x}\nabla_x f_1(H') + f_1 \dot{x}\nabla_x f_2 = N d p_\eta f_1 f_2 = f_1 \dot{x}\nabla_x f_2 \tag{2.25}
$$

where the last equality follows from the fact that H' is a constant of motion. From Eq. 2.25 it follows that:

$$
f_2 \propto e^{dN\eta} \tag{2.26}
$$

The invariant measure is then:

$$
d\mu \propto \delta(H' - C) e^{dN\eta} d\eta dp_\eta \prod_{i=1}^{dN} dq_i dp_i = \delta(H' - C) e^{dN\eta} d\eta dp_\eta d\Gamma \tag{2.27}
$$

The microcanonical partition function of the system coupled to a Nosé-Hoover thermostat takes the form:

$$
\Omega = \int \delta(H' - C) e^{dN\eta} d\eta dp_\eta d\Gamma \tag{2.28}
$$

Now we only have to integrate over the thermostat's degrees of freedom to obtain the partition function of the system. We can eliminate the δ measure integrating over $d\eta$. From Eq. 2.22 it follows:

$$
\eta = (C - H - p_\eta^2/2Q)/Lk_bT = \beta(C - H - p_\eta^2/2Q)/L \tag{2.29}
$$

Substituting into Eq. 2.28 and integrating over dp_η we obtain:

$$\Omega \propto \int e^{-\beta dNH/L} d\Gamma \tag{2.30}$$

If we choose $L = dN$ the partition function of the system (i.e. after integrating the thermostat degrees of freedom) is canonical. If $\sum_i F_i = 0$ system 2.21 has an additional constant of motion:

$$(\sum_i p_i)e^\eta = constant \tag{2.31}$$

In this case the problem is solved with a chain of thermostats which eliminate this symmetry (see [64]). Equations of motion, for a chain of M thermostats, write:

$$\begin{aligned}
\dot{r}_i &= \frac{p_i}{m_i} \\
\dot{p}_i &= F_i - \frac{p_{\eta_1}}{Q_1} p_i \\
\dot{\eta}_i &= \frac{p_{\eta_i}}{Q_i} \quad i = 1, ..., M \\
\dot{p}_{\eta_1} &= \sum_{i=1}^{dN} \frac{p_i^2}{m_i} - dNk_BT - \frac{p_{\eta_2}}{Q_2} p_{\eta_1} && (2.32) \\
\dot{p}_{\eta_k} &= \frac{p_{\eta_{k-1}}^2}{Q_{k-1}} - dNk_BT - \frac{p_{\eta_{k+1}}}{Q_{k+1}} p_{\eta_k} \quad k = 2, ..., M-1 && (2.33) \\
\dot{p}_{\eta_M} &= \frac{p_{\eta_{M-1}}^2}{Q_{M-1}} - dNk_BT && (2.34)
\end{aligned}$$

Let us now consider the partition function of a constrained system. We are interested in the relation between the partition function of a constrained system with the unconstrained one. Consider a system whose equations of motion have the form:

$$\begin{aligned}
\dot{r}_i &= p_i/m_i \\
\dot{p}_i &= F_i
\end{aligned} \tag{2.35}$$

If we add a constraint $\sigma(r_i) = 0$ we have to generate trajectories satisfying:

$$\begin{aligned}
\sigma(r_i) &= 0 \\
\dot{\sigma}(r_i) &= \sum_i p_i/m_i \nabla_i \sigma
\end{aligned} \tag{2.36}$$

Equations of motion become:

$$\begin{aligned} \dot{r}_i &= p_i/m_i \\ \dot{p}_i &= F_i - \lambda \nabla_i \sigma(r_j) \end{aligned} \tag{2.37}$$

where λ is the Lagrange multiplier. With a little algebra, from $\ddot{\sigma}(r_i) = 0$, λ can be obtained as a function of phase-space variables (see [64]):

$$\begin{aligned} \dot{r}_i &= p_i/m_i \\ \dot{p}_i &= F_i - \Big(\frac{\sum_j F_j \cdot \nabla_j \sigma/m_j + \sum_{j,k} (\nabla_j \nabla_k \sigma) \cdot\cdot\, p_j p_k/(m_j m_k)}{\sum_l (\nabla_l \sigma)^2/m_l} \Big) \nabla_i \sigma \end{aligned} \tag{2.38}$$

We can write Eq. 2.5 for this non Hamiltonian dynamical system:

$$\dot{x} \nabla_x f = \frac{2 \sum_i \nabla_i \sigma/m_i \cdot \nabla_i \sum_j \nabla_j \sigma \cdot p_j/m_j}{\sum_j (\nabla_j \sigma)^2/m_j} f \tag{2.39}$$

As for the case of Nosé-Hoover thermostat, because σ and $\dot{\sigma}$ are constant of motion, f must have the form:

$$f = \delta(\sigma)\delta(\dot{\sigma}) f_3(..) \tag{2.40}$$

from which we obtain the equation for f_3:

$$\dot{x} \nabla_x f_3 = \frac{2 \sum_i \nabla_i \sigma/m_i \cdot \nabla_i \sum_j \nabla_j \sigma \cdot p_j/m_j}{\sum_l (\nabla_l \sigma)^2/m_l} f \tag{2.41}$$

whose solution is:

$$f_3 = \sum_l (\nabla_l \sigma)^2/m_l = Z(r) \tag{2.42}$$

It follows that if $f(p_i, r_i)$ is the density distribution function of the unconstrained system, the partition function of the constrained ensemble is:

$$\Omega_\sigma = \int d\Gamma Z(r) \delta(\sigma(r_i)) \delta(\dot{\sigma}(r_i, p_i)) f(p_i, r_i) \tag{2.43}$$

If the system has more than one constraint $\sigma_k, k = 1, ..., M$, it can be shown (what follows in this section can be found in [10, 11, 57] and references therein) that relation 2.43 writes:

$$\Omega_\sigma = \int d\Gamma \det(Z(r)) \prod_{i=1}^{M} \delta(\sigma(r_i)) \prod_{i=1}^{M} \delta(\dot{\sigma}(r_i, p_i)) f(p_i, r_i) \tag{2.44}$$

where:

$$Z_{lm} = \sum_{i=1}^{dN} \frac{1}{m_i} \frac{\partial \sigma_l}{\partial r_i} \cdot \frac{\partial \sigma_m}{\partial r_i} \quad l, m = 1, ..., M \tag{2.45}$$

Now we want to obtain the relation between constrained and unconstrained averages of observables independent of p_i. To obtain this relation we have to integrate over p_i. An expression of $\delta(\dot{\sigma}())$ in function of p_i is needed. Suppose that the system has sufficiently smooth constraints. If this is the case, the system state can be described in generalized coordinates u:

$$\begin{aligned} u &= (u_1, ..., u_{3N}) = (q_1, .., q_{3N-M}, q_{\sigma_1}, ..., q_{\sigma_M}) = (q, q_\sigma) \\ u_i &= u_i(r_1, .., r_{3N}) \end{aligned} \tag{2.46}$$

The notation (q, q_σ) means that generalized coordinates q_σ and velocities $\dot{q}_\sigma$ are zero if constraints are satisfied. With the same notation, for the conjugate momenta:

$$p_u = (p_{q_1}, .., p_{q_{3N-M}}, p_{\sigma_1}, ..., p_{\sigma_M}) = (p_q, p_\sigma) \tag{2.47}$$

p_σ is equal to $\tilde{p}_\sigma = Z^{-1}\dot{\sigma}$ when $\dot{q}_\sigma = 0$. Then f_σ can be written as:

$$f_\sigma = d\Gamma \delta(\sigma(r_i))\delta(p_\sigma - \tilde{p}_\sigma) f(r_i, p_i) \tag{2.48}$$

If f is the canonical density probability function and the Hamiltonian of the system is of the form $H = K(p_i) + V(r_i)$:

$$f(p_i, r_i) = e^{-\beta H(r_i, p_i)} = e^{-\beta K(p_i)} e^{-\beta V(r_i)} \tag{2.49}$$

with a change of coordinates $r \to q$, it can be shown that the reduced density of f_σ is:

$$f_\sigma(r_i) = \int \prod_i dp_i \; \delta(p_\sigma - \tilde{p}_\sigma) f_\sigma(r_i, p_i) = \frac{e^{-\beta V(r_i)} (\det Z)^{1/2} \delta(\sigma(r_i))}{\int \prod_i dr_i \; \{...\}} \tag{2.50}$$

From the last equation it is clear that if O is an observable depending only on r_i then:

$$\langle O(r_i)\delta(\sigma(r_i))\rangle = \langle (\det Z)^{-1/2} O(r_i)\rangle_\sigma \tag{2.51}$$

where $\langle ... \rangle$, $\langle ... \rangle_\sigma$ are respectively averages over unconstrained and constrained trajectories.

2.4 Symplectic integrators

As we have seen, to simulate a system of particles in microcanonical ensemble, we have to solve the Hamiltonian equations of motion:

$$\dot{q}_i = \frac{\partial H}{\partial p_i} = \{p_i, H\}$$
$$\dot{p}_i = -\frac{\partial H}{\partial q_i} = \{q_i, H\}$$

These equations define an evolution $(q_i(0), p_i(0)) \to (q_i(t), p_i(t))$ which is called symplectic.

The attainment of equilibrium starting from a non-equilibrium configuration is however another problem. We have to start the simulation in a state as close as possible to equilibrium.

Once we choose an initial condition, we have to solve in an approximate way the $3N$ equations of motion. Even if, due to Lyapunov instability, numerical errors will produce an exponential divergence from the "true trajectory" for a long time, if the numerical algorithm preserves some characteristic properties of the true trajectory, the approximate trajectories will be always close to some characteristic trajectory. In general the algorithm has to preserve symmetries of the Hamiltonian equations of motion. Hamiltonian equations have N Poincaré constants of motion [18, 22] which are:

$$\omega^{2k} = \sum_{1 \leqslant i_1 \leqslant ... \leqslant i_k} dp^{i_1} \otimes ... \otimes dp^{i_k} \otimes dq_{i_1} \otimes ... \otimes dq_{i_k} \tag{2.52}$$

Moreover Hamiltonian equations are time-reversible and conserve the total energy H. An algorithm which conserves differential forms 2.52 and is time reversible is called a symplectic integrator.

Consider an algorithm of type:

$$\begin{aligned} do \;\; k = 1, ..., N \\ \underline{p}^k = \underline{p}^{k-1} + dt\, \underline{\alpha}^k F(\underline{q}^{k-1}) \\ \underline{q}^k = \underline{q}^{k-1} + dt\, \underline{\beta}^k K(\underline{p}^k) \end{aligned} \tag{2.53}$$

where $\underline{p} = (p_1, .., p_{3N}), \underline{q} = (q_1, .., q_{3N})$, $F(\underline{q}) = -\partial V(\underline{q})/\partial \underline{q}$, $K(\underline{p}) = \partial T(\underline{p})/\partial \underline{p}$ ($V(\underline{q}), T(\underline{p})$ are respectively the potential and the kinetic energy of the system), $\underline{\alpha}^k, \underline{\beta}^k$ are numerical constants and dt the time step. It is easy to see

that algorithm 2.53 consists of two steps:
an exact evolution under a Hamiltonian $V(\underline{q})$ for a time $\underline{\alpha}^k dt$;
an exact evolution under a Hamiltonian $T(\underline{p})$ for a time $\underline{\beta}^k dt$;
Algorithm 2.53 is a succession of two symplectic transformations (since Hamiltonian) which is a symplectic transformation and thus it is symplectic [25].

A way to obtain symplectic stable integrators is to use the Liouvillian formalism. In this formalism it is easy to obtain that total energy H is approximately conserved uniformly in time and the time-reversibility algorithm is obvious (see next). Consider an arbitrary function $g(q_i, p_i)$. If (q_i, p_i), briefly (q,p), evolves in some way, the evolution of g can be formally written as:

$$\dot{g}(q,p) = iL\, g(q,p) \;\; with \;\; L = \dot{r}\frac{\partial}{\partial r} + \dot{p}\frac{\partial}{\partial p} = L_r + L_p \tag{2.54}$$

whose formal solution is:

$$g[q(t), p(t)] = e^{iLt} g[q(0), p(0)] \tag{2.55}$$

This solution is formal because, since L_r and L_p do not commute, it is not possible to evaluate the action of e^{iLt} on an arbitrary function of (q, p). However it is possible to evaluate the action of e^{itL_r} or e^{itL_p}, on an arbitrary function g (see [20]):

$$\begin{aligned} e^{itL_r} g[q(0), p(0)] &= g[q(0) + \dot{r}(0)\, t\, ,\, p(0)] \\ e^{itL_p} g[q(0), p(0)] &= g[q(0)\, ,\, p(0) + \dot{p}(0)\, t] \end{aligned} \tag{2.56}$$

Since L_r and L_p do not commute:

$$e^{it(L_r+L_p)} \neq e^{itL_r} e^{itL_p} \tag{2.57}$$

At this point we can use the Trotter expansion:

$$e^{(A+B)} = \lim_{k\to\infty} \left(e^{A/2k} e^{B/k} e^{A/2k}\right)^k = \left(e^{A/2k} e^{B/k} e^{A/2k}\right)^k e^{O(1/k^2)} \tag{2.58}$$

Let us make the position:

$$\begin{aligned} \frac{A}{2k} &= iL_p \frac{t}{2k} = iL_p \frac{dt}{2} \\ \frac{B}{k} &= iL_r \frac{t}{t/k} = iL_r dt \end{aligned} \tag{2.59}$$

Using Eq. 2.56 it is easy to verify that (see [20]):

$$e^{iL_p dt/2} e^{iL_r dt} e^{iL_p dt/2} g[(q(0), p(0)] = g[p(dt), q(dt)] \tag{2.60}$$

where:

$$\begin{aligned} p(dt/2) &= p(0) + \dot{p}(0)dt/2 \\ q(dt) &= q(0) + p(dt/2)dt \\ p(dt) &= p(dt/2) + \dot{p}(dt/2)dt/2 \end{aligned} \tag{2.61}$$

If the system is Hamiltonian $\dot{p} = -\frac{\partial V}{\partial r} = F$ and $\dot{r} = \frac{p}{m}$ then Eq. 2.61 writes:

$$\begin{aligned} q(dt) &= q(0) + \dot{r}(0)dt + \frac{1}{2}F(0)\frac{dt^2}{m} \\ \dot{r}(dt) &= \dot{r}(0) + \frac{1}{2}\Big(F(0) + F(dt)\Big)\frac{dt}{m} \end{aligned} \tag{2.62}$$

If we want to obtain velocities and positions at time t, we have to repeat algorithm 2.62 k times. This is the so called velocity-Verlet algorithm. Eq .2.61 are of type of Eq. 2.53, so the velocity-Verlet algorithm is symplectic.

Let us rewrite Eq. 2.58 with position 2.59:

$$e^{it(L_r+L_p)} = e^{O(t/k^2)}(e^{itL_p/2k} e^{itL_r/k} e^{itL_p/2k})^k = e^{O(dt^2)}(e^{idtL_p/2} e^{idtL_r} e^{idtL_p/2})^k \tag{2.63}$$

Consider the action of operator 2.63 on a generic constant of motion $C[q, p]$:

$$\begin{aligned} e^{it(L_r+L_p)} C[q(0), p(0)] &= C[q(0), p(0)] = \\ = e^{O(dt^2)}(e^{idtL_p/2} e^{idtL_r} e^{idtL_p/2})^k C[q(0), p(0)] \end{aligned} \tag{2.64}$$

Clearly:

$$(e^{idtL_p/2} e^{idtL_r} e^{idtL_p/2})^k C[q(0), p(0)] \tag{2.65}$$

represents the value of the constant of motion C when phase space coordinates are evolved under the approximate propagator:

$$e^{iL_{ps}t} = (e^{idtL_p/2} e^{idtL_r} e^{idtL_p/2})^k \tag{2.66}$$

let us indicate these trajectories as (q_{ps}, p_{ps}) ("ps" means pseudo-trajectories), then Eq. 2.64 writes:

$$C[q(0), p(0)] \simeq C[q_{ps}(t), p_{ps}(t)] + O(dt^2) \tag{2.67}$$

this equation explains why no energy drift is present in algorithms derived with this formalism. (for a rigorous derivation see [5, 26, 53]). Obviously time reversibility follows from:

$$e^{iL_{ps}t}e^{-iL_{ps}t} = 1 \tag{2.68}$$

The Velocity-Verlet algorithm was derived assuming Hamiltonian equations of motion and with a specific factorization of e^{iLt} which is arbitrary. In general external terms in factorization of the 2.63 type are those which contains higher frequencies as for example in multiple time step algorithms [20].

For non-Hamiltonian time-evolution the same type of formalism can be used to derive time-reversible algorithms [62]. In fact property 2.67 remains valid also for non-Hamiltonian flows. This assures that, for any time, approximated trajectories will not produce a drift in constants of motion. In the case of constrained systems, a special type of symplectic algorithm, called SHAKE, is used to satisfy constraints (see [3, 37, 55]).

2.5 The SHAKE algorithm

SHAKE is an algorithm to deal with constrained systems (see [55]). Let us suppose to have a N particle system subject to a constraint σ:

$$\sigma(r_1, ..., r_{3n}) = 0 \tag{2.69}$$

If we indicate with $\lambda(t, r_1, ..., r_{3N}, \dot{r}_1, ..., \dot{r}_{3N})$ the Lagrange multiplier at time t (see Eq. 2.38) we should try to obtain a solution of equations of motion simply by writing:

$$r_i(h) = r_i(0) + .. - \frac{1}{2}\frac{h^2}{m_i}\lambda(0, ..)\nabla_{r_i}\sigma(r_1, ..., r_{3N})(0) \tag{2.70}$$

where ... stays for some symplectic integrator and h is the time step. If one solves equations of motion calculating at each time step the constrained force and adds this force to the internal force as in Eq. 2.70, constraints are violated more and more during the motion. In fact, the constrained force is not able to bring the system at $\sigma = 0$. then, in this way, only the constrained force compatible with $\sigma(r^{un}) \neq 0$ is calculated. It is clear now that this is not a good method to solve equations of motion of systems with constraints.

The central idea of the SHAKE algorithm is to use constraints to obtain a λ which exactly satisfies $\sigma = 0$. Let us indicate with $..^c, ..^{un}$ the constrained and unconstrained quantities. At each time step:

1) starting from $(r_i^c(0), \dot{r}_i^c(0))$ one solves unconstrained equations of motion with some algorithm to obtain $(r_i^{un}(h), \dot{r}_i^{un}(h))$.
2) Substituting the following expression of r_i^{un}:

$$r_i^c(h) = r_i^{un}(h) - \frac{h^2}{m_i}\lambda\nabla_{r_i}\sigma(r_1, ..., r_{3N})_{|r^{un}(0)} = r_i^{un}(h) - \frac{h^2}{m_i}\lambda(\nabla_i\sigma)(0) \tag{2.71}$$

in Eq. 2.69:

$$\sigma(r_1^c(h), ..., r_{3N}^c(h)) = 0 \tag{2.72}$$

λ is calculated as to satisfy exactly the constraint. Unfortunately only in some cases λ can be calculated exactly and anyway the analytical calculation of λ

should result computationally time consuming. For these reasons the following iterative procedure is adopted to calculate λ. Let us write:

$$\begin{aligned} \sigma(h) &= \sigma(r_i^{un}(h)) \\ (\nabla_{r_i}\sigma)(r_1^{un}(h), ..., r_{3N}^{un}(h)) &= \nabla_i\sigma(h), \quad (\nabla_{r_i}\sigma)(r_i^{un}(0)) = \nabla_i\sigma(0) \end{aligned} \tag{2.73}$$

Using Eq. 2.71, Eq. 2.72 can be approximatively rewritten as:

$$\begin{aligned} 0 = \sigma(h) &= \sigma\Big(r^c(0)\Big) + \sum_{i=1}^{3N} \nabla_i\sigma(h)[r_i^c(h) - r_i^{un}(h)] + O(h^4) = \\ \sigma(h) &= 0 + \sum_{i=1}^{3N} \frac{h^2}{m_i}\nabla_i\sigma(h)\nabla_i\sigma(0)\lambda + O(h^4) \rightarrow \\ \rightarrow \lambda h^2 &\simeq \frac{\sigma(h)}{\sum_{i=1}^{3N}\frac{1}{m_i}\nabla_i\sigma(h)\nabla_i\sigma(0)} \end{aligned} \tag{2.74}$$

Obviously, since a Taylor expansion has been used, with the obtained λ results $\sigma(r_i^c(h)) \neq 0$. Then the procedure has to be repeated until the desired convergence is reached.

In presence of more than one constraint $\sigma_1, ..., \sigma_l$ Eq. 2.71, 2.74 become:

$$r_i^c(h) = r_i^{un}(h) - \sum_{k=1}^{l} \frac{h^2}{m_i}\lambda_k(\nabla_i\sigma_k)(0) \tag{2.75}$$

$$\sigma_k(h) = \sum_{i=1}^{3N} \frac{h^2}{m_i} \sum_{k'=1}^{l} \nabla_i\sigma_k(h)\nabla_i\sigma_{k'}(0)\lambda_{k'} \tag{2.76}$$

In this case a matrix inversion is needed, which is computationally expensive. For this reason a modification of the previous scheme is applied, which considers the constraints in succession. Instead of Eq. 2.76 one writes:

$$r_i^c(h) \simeq r_i^{un} - \frac{h^2}{m_i}\lambda_k\nabla_i\sigma(0) \tag{2.77}$$

and hence for λ_k the obtained estimation is:

$$\lambda_k h^2 = \frac{\sigma_k(0)}{\sum_{i=1}^{3N}\frac{1}{m_i}\nabla_i\sigma_k(h)\nabla_i\sigma_k(0)} \tag{2.78}$$

This is repeated for each constraint. The entire cycle is repeated until the desired convergence on each constraint is reached.

The SHAKE algorithm cannot be symplectic in the strictly sense. In fact the condition:

$$\sum_i \dot{r}_i \nabla_i \sigma = 0 \tag{2.79}$$

which represents the orthogonality of velocities to constrained force is not guaranteed. A combination of SHAKE algorithm with the velocity-Verlet algorithm can be used to ensure this condition. Such generalized algorithm is called RATTLE [3]. However it can be shown that SHAKE (if a Verlet scheme is used for unconstrained dynamics) and RATTLE algorithms are equivalent at $h/2$ (see velocity-Verlet algorithm, Eq. 2.61) and the last one is symplectic [37].

2.6 A SHAKE-like predictor-corrector algorithm

In this section, to fix ideas, we always refer to NVE type equations of motion. In general symplectic integrators, like Verlet or velocity-Verlet, violate in a certain sense the second Newton law. In fact the general scheme to obtain solutions of equations of motion through numerical integrators have the following characteristic: consider a single step evolution: $(r_i(0), v_i(0))$ are transformed in some way to $(r_i(h), v_i(h))$. For $t \in [0, h[$ accelerations are not equal to $f_i(t)/m_i$, but are equal to $f_i(0)/m_i$ and then the second Newton law is not strictly verified. At time h it is placed $a_i(h) = f_i(h)/m_i$. There are some types of algorithms, called predictor-corrector, which try to overcome this problem [2, 20]. These schemes generate more correct solutions for short times, but produce a systematic energy drift since they are not symplectic. No symplectic property implies that they are also not time-reversible. Then, even if such a scheme permits a longer time step, they are usually not used because a single time-step is computationally more expansive than a symplectic integrator time-step (forces have to be calculated more times for each step) and moreover they don't have good long time properties (energy drift, time reversibility).

In general every type of equation of motion has some conserved quantity. NVE equations of motion conserve the total energy of the system:

$$H = \sum_{i=1}^{3N} \frac{1}{2} m \dot{r}_i^2 + U(r_i) = constant \tag{2.80}$$

Since numerical schemes as we have just said violate the second Newton law, one should impose the constraint of total energy conservation. In other words one can use some symplectic scheme to solve unconstrained equations and at each step impose the total energy constraint in a way similar to SHAKE algorithm. The case treated in the previous section refers to holonomic constraints that are constraints which don't depend on particles velocities. The Gauss principle of least constraint permits to deal also with not holonomic constraints when the constraints are quadratic in particles velocities. This principle is also used to write isokinetic equations of motion [15]. In [17] it is given the expression of force due to the presence of a not holonomic constraint. Equations of motion can be written in general as:

$$m\ddot{a}_i = f_i + R_i \tag{2.81}$$

The reaction forces R_i will depend in general on the type of constraint. In the case of total energy constraint, the Newton second law has to be obtained, then $R_i = 0$. This fact is true for every constraint which is a constant of motion. However, since a numerical algorithm is used, these constraint forces are not equal to zero. As in the case of holonomic constraints, Gauss's principle of least constraint doesn't help us since, even if R_i were calculated, these forces are unable able to bring the system at the initial value of H. Also in this case a SHAKE-like scheme can be used. Let us indicate the $\lim_{t\to h^-} a(t)$ with $a(h^-)$. The following SHAKE-like algorithm can be used to impose total energy conservation at each step (notation are as in Eq. 2.73):

1) starting from $(r_i^c(0), \dot{r}_i^c(0))$ one solves unconstraint equations of motion with some algorithm to obtain $(r_i^{un}(h), \dot{r}_i^{un}(h))$.
2) Substituting the following expressions of $r_i^{un}, \dot{r}_i^{un}$:

$$\begin{aligned} r_i^c &= r_i^{un}(h) + \frac{1}{2}\lambda\frac{h^2}{m}R_i(h^-) = r_i^{un}(h) + \frac{1}{2}\lambda\frac{h^2}{m}[f_i(h^-) - ma_i(h^-)] \\ \dot{r}_i^c &= \dot{r}_i^{un}(h) + \lambda h R_i(h^-) = \dot{r}_i^{un}(h) + \lambda h[f_i(h^-) - ma_i(h^-)] \end{aligned} \tag{2.82}$$

in Eq. 2.80 and expanding the equation to order $O(\lambda^2 h^2)$, λ is calculated. The cycle is repeated until the desired convergence is reached. In the case of particles of equal mass, it is easy to verify that the obtained expression of λ is:

$$\begin{aligned} \lambda &= \frac{-b + \sqrt{b^2 - ac}}{a} \quad if \quad b \geqslant 0 \\ \lambda &= \frac{-b - \sqrt{b^2 - ac}}{a} \quad if \quad b < 0 \end{aligned} \tag{2.83}$$

where:

$$\begin{aligned} a &= mh\sum_i [f_i(h^-) - ma_i(h^-)][f_i(h^-) - ma_i(h^-)] \\ b &= m\sum_i \dot{r}_i^{un} h[f_i(h^-) - ma_i(h^-)] - \frac{h}{2}\sum_i f_i(h^-)[f_i(h^-) - ma_i(h^-)] \\ c &= \frac{-2\Delta E}{h}; \quad \Delta E = H(0) - H(r_i^{un}, \dot{r}_i^{un}) \end{aligned} \tag{2.84}$$

This integration scheme has exactly the same form as a second order predictor-corrector scheme. The only differences is that in a predictor-corrector algorithm in Eq. 2.82, instead of λ there are numerical constants [2, 20]. Symplectic

properties of this algorithm I suppose can be obtained in a way similar to[37] where the SHAKE algorithm is treated. This algorithm was introduced in a classical molecular dynamics code. The case studied consists of a system of Si atoms interacting through a classical Stillinger-Weber potential. The initial temperature was of 900 K. In Fig. 2.1 the behaviour of pressure and total energy of a 64 atoms system with a time step length of 5 fs are reported. From these figures it is clear that typical fluctuations of pressure are equal for both algorithms but the total energy fluctuations are practically zero for this SHAKE like algorithm. For potential and kinetic energy the behaviour is the same as for pressure. The numerical precision of this algorithm was tested calculating the divergence (see App.A.2) of two trajectories differing for the initial position of a Si atom which was shifted of 0.1 Bohr in each direction. In Fig. 2.2 a comparison of a velocity-Verlet algorithm with this algorithm is reported and it is clear that the asymptotic slope (Lyapunov exponents depend on this slope) is better estimated with this method when one increases the time-step length. The code performance is reported in Tab. 2.1.

Algorithm	Time step (fs)	ΔE (eV/Atom)	CPU time/step (ms)
Velocity-Verlet	1	1.6E-4	7
Velocity-Verlet	5	2.2E-3	8
This algorithm	1	$<$ 10E-6	11
This algorithm	5	$<$ 10E-6	16
This algorithm	7	$<$ 10E-6	28

Table 2.1: Code performance table for 1000 Si atoms with a 1.6 GHz G5.

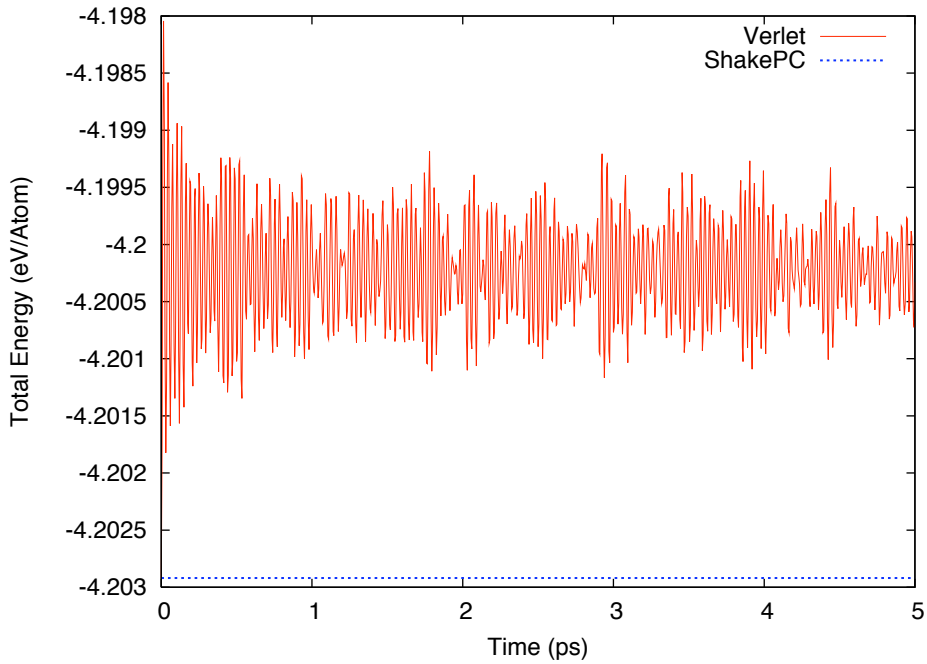
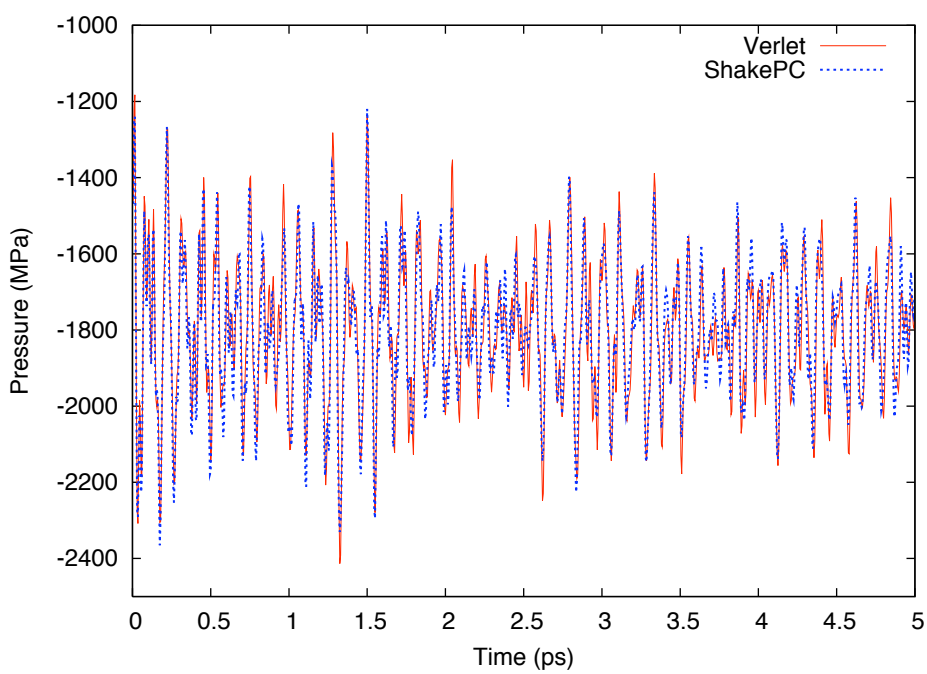


Figure 2.1: Comparison of velocity-Verlet algorithm with this algorithm for $h = 5$ fs. Up total energy (the initial total energies are equal). Down the pressure.



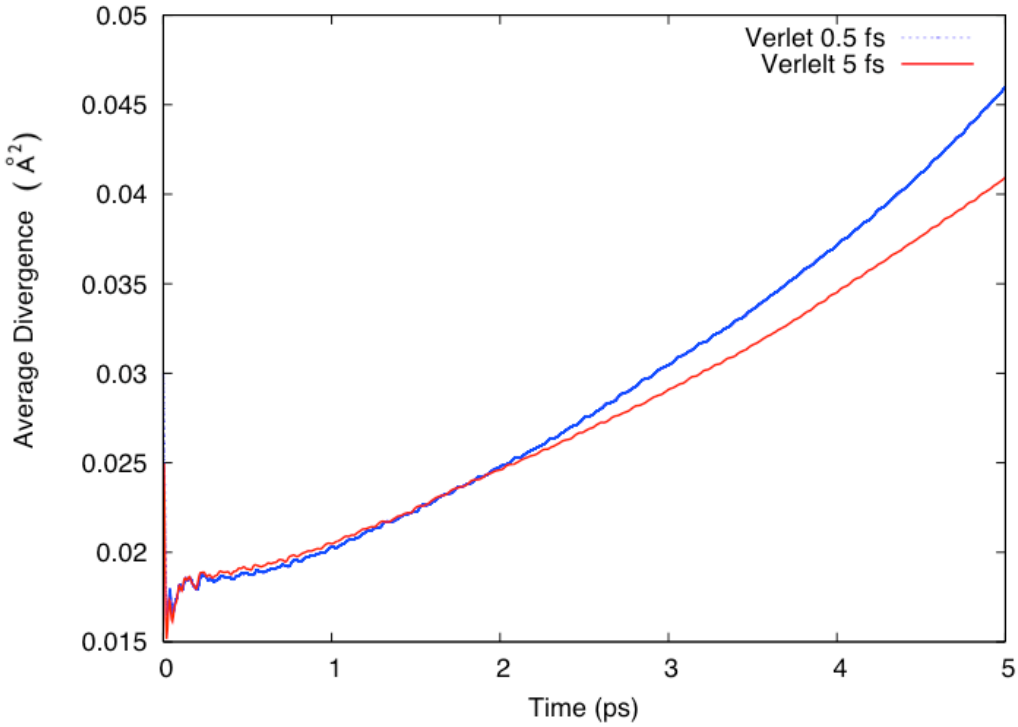
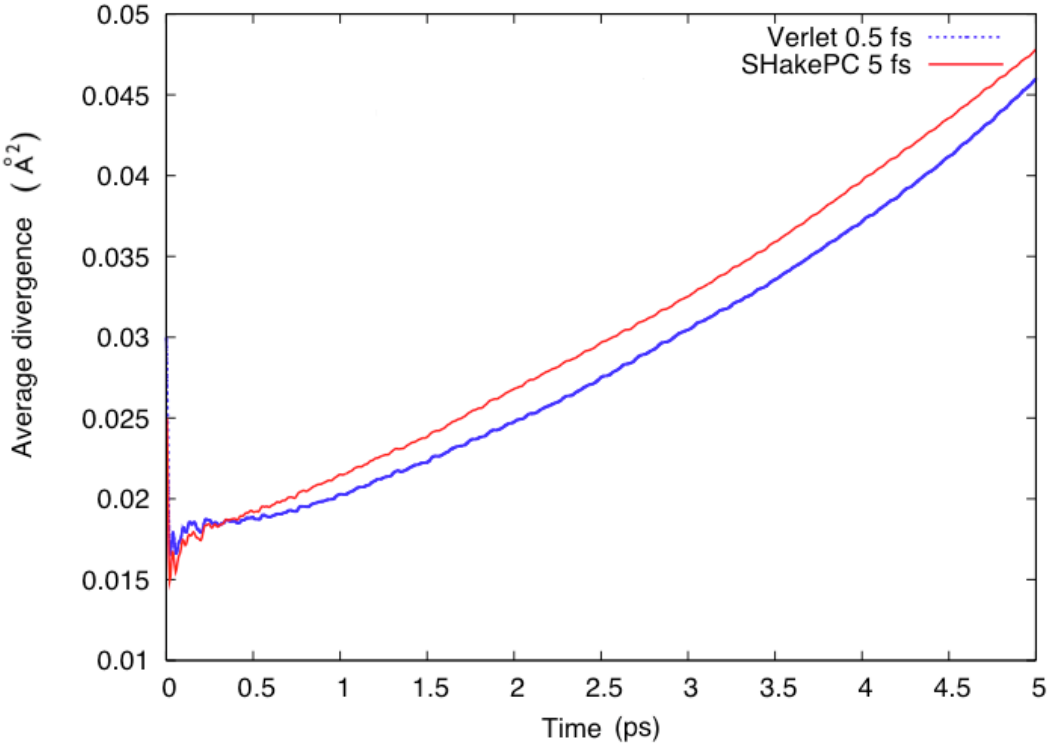


Figure 2.2: Average divergence (defined in A.2) comparison.



2.7 Free energies from MD

To know which state, let us say ξ_1, ξ_2, is stable at given values of (N,V,T), we have to compute the free energy difference between these states:

$$F(N,V,T) = -k_B T \ln Q(N,V,T) \;\; Q(N,V,T) = \int d\Gamma e^{-\beta H} \tag{2.85}$$

Free energy is the thermodynamic potential in the canonical ensemble. Differences in free energy (as for all potentials) depend only on the states ξ_1, ξ_2. Free energies calculations are used to study the phase-diagrams of substances at given (N,V,T) conditions. However, quantities like Q, which are called thermal quantities, cannot be computed directly in a MD simulation because they are not canonical averages over phase space but depend on the phase space volume accessible to the system. The same situation is present in real experiments [20]. Let us introduce the reversible work W (the free energy difference) to bring the system from a state ξ to ξ' (what follows in this section can be found in [10, 11, 57] and references therein). From Eq. 2.85 it follows that:

$$W(\xi') = -k_B T \ln P_\xi(\xi') \;\; P_\xi(\xi') = \frac{1}{Q} \int d\Gamma e^{-\beta H} \delta(\xi(r_i) - \xi') \tag{2.86}$$

From Eq. 2.86 we can write :

$$\begin{aligned} W(\xi_2) - W(\xi_1) &= \int_{\xi_1}^{\xi_2} d\xi' \frac{dW(\xi')}{d\xi'} = \int d\xi' \frac{\langle -\frac{\partial H}{\partial \xi} \delta(\xi(r_i) - \xi') \rangle}{\langle \delta(\xi_2(r_i) - \xi') \rangle} = \\ &= \int_{\xi_1}^{\xi_2} d\xi' \Big\langle -\frac{\partial H}{\partial \xi} \Big\rangle_{\xi'}^{cond} \end{aligned} \tag{2.87}$$

where $\langle ... \rangle$ denotes canonical averages. The definition of conditional canonical averages is:

$$\langle ... \rangle_{\xi'}^{cond} = \frac{\langle ... \delta(\xi(r_i) - \xi') \rangle}{\langle \delta(\xi(r_i) - \xi') \rangle} = \frac{\int d\Gamma e^{-\beta H} ... \delta(\xi(r_i) - \xi')}{\int d\Gamma e^{-\beta H} \delta(\xi(r_i) - \xi'))} \tag{2.88}$$

Eq. 2.87 is called a thermodynamic integration. If we want to use this expression to calculate free energy differences in an MD run, we have to do unconstrained equilibrium time averages of $\langle -\frac{\partial H}{\partial \xi} \rangle_{\xi'}^{cond}$ for different values of

ξ' and then calculate the integral 2.88. However if ξ' is "rare" we should have to do a "long" run to have a sufficient statistics for ξ' (for system we studied we are able to simulate not more than some *ps* of dynamics). We could use at this point what has been said at the end of Sec. 2.3, Eq. 2.51. In fact for an observable O which depends only on r_i we have:

$$\langle O(r_i)\rangle_{\xi'}^{cond} = \frac{\langle O(r_i)\delta(\xi(r_i)-\xi')\rangle}{\langle \delta(\xi(r_i)-\xi')\rangle} = \frac{\langle Z^{-1/2}O(r_i)\rangle_{\xi'}}{\langle Z^{-1/2}\rangle_{\xi'}} \tag{2.89}$$

Since $\frac{\partial H}{\partial \xi}$ in general depends on p_i this property cannot be used. We have to integrate analytically Eq. 2.88 (for $\cdots = \frac{\partial H}{\partial \xi}$ without $\dot{\xi} = 0$). The result is:

$$\frac{dW}{d\xi'} = \Big\langle -\frac{\partial H}{\partial \xi}\Big\rangle_{\xi'}^{cond} = \frac{\Big\langle Z^{-1/2}\frac{\partial}{\partial \xi}\Big[V - k_B T \ln(|J|Z^{1/2})\Big]\Big\rangle_{\xi'}}{\langle Z^{-1/2}\rangle_{\xi'}} \tag{2.90}$$

To use this result one has to make the coordinate transformation $r \to u$. Moreover the knowledge of J and its partial derivative with respect to ξ is needed. If we call λ the Lagrange multiplier which determines the constraint force, it can be shown that Eq. 2.90 can be rewritten as:

$$\frac{dW}{d\xi'} = \frac{\langle Z^{-1/2}[\lambda + k_B T G]\rangle_{\xi'}}{\langle Z^{-1/2}\rangle_{\xi'}} \tag{2.91}$$

where:

$$G = \frac{1}{Z^2}\sum_{i,j}\frac{1}{m_i m_j}\frac{\partial \xi}{\partial r_i}\cdot\frac{\partial^2 \xi}{\partial r_i r_j}\cdot\frac{\partial \xi}{\partial r_j} \tag{2.92}$$

Moreover λ is independent of the coordinates system. In our calculation ξ is of type $|r_i - r_j| = 0$ for some i, j. In this case Z is constant and $G = 0$.

A generalization of this techniques permits to calculate also time-dependent properties from constrained MD [10, 11].

The method exposed here is an equilibrium method, that is, time averages are calculated in thermodynamic equilibrium. Other techniques use non-equilibrium simulations [20, 54, 69].

3 Introduction to Density Functional Theory

In the previous sections we referred in particular to classical statistical mechanics. In fact, in problems we are interested in, ions can be treated as classical objects. This approximation is valid if:

$$\lambda_T \ll a \tag{3.1}$$

where a is the average distance between ions and $\lambda_T = \sqrt{\frac{2\pi h^2}{MKT}}$ the thermal de Broglie wavelength. The main problem, which we don't take into account, concerns interatomic forces calculation. In some cases, forces can be obtained using effective classical potentials. In other words the system properties are calculated solving classical equations of motion deriving from a classical Hamiltonian of the form:

$$H(r_i, p_i) = \sum_i \frac{p_i^2}{2M_i} + \sum_{i<j} u^2(r_i, r_j) + \sum_{i<j<k} u^3(r_i, r_j, r_k) + ... \tag{3.2}$$

In other cases forces have to be calculated from electronic configurations, that is from "ab-initio" principles. In this case the quantum characteristic of electrons is taken into account but only partially because, due to the small electronic mass, electronic motion time-scale is lower than atomic one. This is called Born-Oppenheimer approximation, which however is valid in a wide range of interesting cases in condensed matter physics. Under this assumption one can uncouple the equations determining the electronic eigenfunctions from those determining the atomic ones.

3.1 Born-Oppenheimer approximation

To see which assumptions are done in Born-Oppenheimer approximation (see [8]), let us write the time-independent Schrodinger equation for a full quantum mechanical system:

$$[K_e + K_I + V_{Ie}(R_I, r_i) + V_{ee}(r_i) + V_{II}(R_I)]\psi(r_i, R_i) = E\psi(R_I, r_i) \tag{3.3}$$

Where e, I stay respectively for electrons and ions, K is the kinetic energy and V the interaction potential. This equation, written for electrons at fixed ions positions, reduces to:

$$[K_e + V_{Ie}(r_i, R_I) + V_{ee}(r_i)]\phi_p(r_i|R_I) = E_p(R_I)\phi_p(r_i|R_I) \tag{3.4}$$

In the last equation, ions positions are parameters of electronic wavefunction. We can choose eigenfunctions $\phi_p(r_i|R_I)$ as complete set of eigenfunctions in which expand $\psi(R_I, r_i)$ which is the full solution of Eq. 3.3, that is:

$$\psi(R_I, r_i) = \sum_p c_p(R_I)\phi_p(r_i|R_I) \tag{3.5}$$

The coefficients $c_p(R_I)$ represent ionic wavefunctions if the electronic state is q and are found inserting Eq. 3.5 in Eq. 3.3 and projecting it on a generic $\phi_q(R_I|r_i)$:

$$\sum_p [\int dr_1....dr_n \phi_q^* K_I \phi_p c_p(R_I)] + E_q(R_I)c_q(R_I) = Ec_q(R_I) \tag{3.6}$$

where:

$$\begin{aligned} K_I\phi_p c_p(R_I) \quad &= -\tfrac{1}{2M_I}\textstyle\sum_I [c_p(R_I)\nabla_I^2\phi_p(r_i|R_I) + 2\nabla_I c_p(R_I) * \nabla_I\phi_p(r_i|R_I) + \\ &\quad + \phi_p(r_i|R_I)\nabla_I^2 c_p(R_I)] \end{aligned} \tag{3.7}$$

The Born-Oppenheimer approximation consists on the following assumptions:

$$|\nabla_I\phi_q(R_I|r_i)| << |\nabla_I c_q(R_I)| \quad for\, all\, q, I \tag{3.8}$$

Eq. 3.8 can be interpreted as the condition that ionic wavefunctions are much more localized than electronic ones. In this case equations Eq. 3.6 uncouple and $c_q(R)$ satisfy:

$$[\sum_I K_I + E_q(R_I)]c_q(R_I) = Ec_q(R_I) \tag{3.9}$$

This equation means that ions wavefunctions (in the state q) are obtained solving the Schrodinger equation in the effective potential E_q.

Due to their small mass, the Fermi temperature of electrons is very high so that, in many situations, they can be considered to be in ground state. If the gap between the ground and the excited states remains large during atomic motion, electrons remain in their ground state also during atomic motion.

In conclusion for classical ions, using the Hellmann-Feynman theorem, the whole problem can be reduced to the following form:

$$M_I \ddot{R}_I = -\frac{\partial E(R_I)}{\partial R_I} - \frac{dV_{II}(R_I)}{dR_I} \tag{3.10}$$

where:

$$E(R_I) = \min_{\psi} E(\psi, R_I) = \min_{\psi} \langle \psi | K_i + V_{ee} + V_{eI} | \psi \rangle \tag{3.11}$$

Although many approximations have been done until this point, the previous equations are still computationally very expansive. The reason is that, for every ions configuration we have to solve Eq. 3.11 for an antisymmetric (the problem grows like $N!$) function $\psi(r_1, ..., r_n)$ which depends on $3N$ electronic positions: this is very expansive even for a small number of electrons.

3.2 Hoenberg-Kohn theorem

To reduce the computationally cost of simulations "Density Functional Theory" is used to treat the electronic subsystem. DFT is a reduction theory which simplifies the problem of previous section written in Eq. 3.10. Given a wavefunction $\psi(r_1, .., r_n)$ of a system of N electrons, its density is defined as:

$$n(r) = N \int d^3r_2...d^3r_n |\psi(r, r_2, .., r_n)|^2 \tag{3.12}$$

in this way:

$$\int d^3r\, n(r) = N \tag{3.13}$$

DFT theory shows that electronic ground state can be uniquely described with the scalar function n.

The Hohenberg-Kohn theorem [27] has represented the starting point of the theory but only with the Kohn-Sham method [34] the theory became of real applicability. Let us begin with the exposition of Hohenberg-Kohn theorem and related variational principle. In what follows we consider an Hamiltonian of the form:

$$H = K_e + V_{ee} + V_{ext} = K_e + V_{ee} + V_{Ie} \tag{3.14}$$

The Hohenberg-Kohn theorem states that the correspondence between:

$$V_{ext}(r) \quad \rightarrow \quad \psi_0(r_1, ..., r_n) \quad \rightarrow \quad n(r) \tag{3.15}$$

is bijective (ψ_0 is the ground state function). Thus all properties of electronic ground state system are unique functionals of $n(r)$. In particular the total energy can be expressed as:

$$E[n] = \int d^3rn(r)v_{ext}(r) + K[n] + V_{ee}[n] = \int d^3rn(r)v_{ext}(r) + F_{HK}[n] \tag{3.16}$$

where:

$$K[n] = \langle \psi[n]|\hat{K}|\psi[n]\rangle \tag{3.17}$$

$$V_{ee}[n] = \langle \psi[n]|\hat{V}|\psi[n]\rangle \tag{3.18}$$

Moreover it can be shown that $E_{v_{ext}}[n] = \int d^3 r n(r) v_{ext}(r) + F_{HK}[n]$ is minimal for $n = n_0$, the ground state density of v_{ext}. In fact for each $\psi[n]$ which does not correspond to the ground state wavefunction $\psi[n_0]$ we have:

$$\langle \psi[n] | \hat{K} + \hat{V}_{ee} + \hat{V}_{ext} | \psi[n] \rangle > \langle \psi[n_0] || \hat{K} + \hat{V}_{ee} + \hat{V}_{ext} | \psi[n_0] \rangle \tag{3.19}$$

The functional $F_{HK}[n]$ is not known so that other approximations are needed.

3.3 Kohn-Sham method

Let us rewrite the total energy functional in the following way:

$$E_v[n] = K_s[n] + \int d^3r\, n(r)v(r) + J[n] + E_{xc}[n] \tag{3.20}$$

with:

$$K_s[n] = \sum_i^N \langle\phi_i| - \frac{1}{2}\nabla^2|\phi_i\rangle \tag{3.21}$$

$K_s[n]$ is the kinetic energy of a gas of non-interacting electrons and $J[n]$ represents the hartree mean field term:

$$E_H[n] = J[n] = \frac{1}{2}\int\int d^3r d^3r' \frac{n(r)n(r')}{|r-r'|} \tag{3.22}$$

Finally the exchange and correlation term E_{xc} has the following form:

$$E_{xc}[n] = (K[n] - K_s[n]) + (V_{ee}[n] - J[n]) \tag{3.23}$$

E_{xc} is calculated in various approximation (GGA, LDA,...) with other techniques like quantum Monte Carlo. Now if $n(r)$ minimizes $E_v[n]$ from the Hohenberg-Kohn theorem we know that $n(r)$ also minimizes the energy $E_{v_s}[n]$ of a system of N non interacting electrons in an effective potential v_s:

$$E_{v_s}[n] = \sum_{i=1}^N \langle\phi_i| - \frac{1}{2}\nabla^2 + v_s|\phi_i\rangle \tag{3.24}$$

In fact Hohenberg-Kohn theorem asserts that there is a bijective correspondence between n and v_{ext} which in this case we call v_s. Imposing the minimum condition $\delta E_v = \delta E_{v_s} = 0$ (with $\int d^3r\delta n(r) = 0$) we can obtain an expression for the effective potential v_s. Let us first write:

$$\delta E_v = \delta K_s + \int d^3r\Big(v(r) + \frac{\partial J}{\partial n(r)} + \frac{\partial E_{xc}}{\partial n(r)}\Big)\delta n(r) = 0 \tag{3.25}$$

with:

$$\frac{\partial J}{\partial n(r)} = v_H(r) = \int d^3r' \frac{n(r')}{|r-r'|} \tag{3.26}$$

$$\frac{\partial E_{xc}}{\partial n(r)} = v_{xc}(r) \tag{3.27}$$

for δE_{v_s} we obtain:

$$\delta E_{v_s} = \delta K_s + \int d^3r\, v_s(r)\delta n(r') = 0 \tag{3.28}$$

Substituting δK_s obtained from Eq. 3.32 in Eq. 3.29 we obtain:

$$\delta E_v = \delta K_s + \int d^3r \Big(v(r) + \frac{\partial J}{\partial n(r)} + \frac{\partial E_{xc}}{\partial n(r)} \Big) \delta n(r) = 0 \tag{3.29}$$

with:

$$\frac{\partial J}{\partial n(r)} = v_H(r) = \int d^3r' \frac{n(r')}{|r - r'|} \tag{3.30}$$

$$\frac{\partial E_{xc}}{\partial n(r)} = v_{xc}(r) \tag{3.31}$$

for δE_{v_s} we obtain:

$$\delta E_{v_s} = \delta K_s + \int d^3r\, v_s(r)\delta n(r') = 0 \tag{3.32}$$

Substituting δT_s obtained from Eq. 3.32 in Eq. 3.29 we obtain:

$$v_s(r) = v(r) + \int d^3r' \frac{n(r)}{|r - r'|} + v_{xc}(r, n(r)) \tag{3.33}$$

This is the desired equation for the effective interaction potential v_s. To resolve the KS equations diagonalization techniques or direct minimization techniques can be used. In the first case the method consist of the following steps:

1) with a trial $n(r)$ calculate the KS potential from Eq. 3.33;
2) diagonalize H_{KS}:

$$[-\frac{1}{2}\nabla^2 + v_s(r)]\phi_i'(r) = \epsilon_i'\phi_i'(r) \tag{3.34}$$

and from $\phi_i'(r)$ calculate the new $n'(r)$:

$$n'(r) = \sum_{i=1}^{N} |\phi_i'(r)|^2 \tag{3.35}$$

3) iterate the procedure until $|n(r) - n'(r)|$ is smaller than a chosen value.

Direct minimization techniques (see [51, 58]) are more robust with respect to diagonalization techniques even if more expansive.

3.4 Periodic boundary conditions: plane waves and pseudopotentials

In order to simulate the bulk properties of a physical system, periodical boundary conditions are used. In fact surface effects should produce a large number of simulated atoms near the surface of volume of the considered system. The surface area decays like $N^{-1/3}$ and then for small systems the surface atoms should be a big fraction of the total system. The only use of periodic boundary conditions is not directly a solution of the problem because one should have to consider an infinite numbers of periodic image atoms (or image cells) to calculate forces on the system. Depending on the system size and on the interaction between atoms, to take into account only a finite numbers of image-cells is a good approximation. If the system is sufficiently large or the interaction-range sufficiently short, one can take into account only some image cells (often only the first-image cell) to calculate forces. Moreover a general technique, called Ewald summation (see for example [20]), can be used to calculate the corrections of the interaction truncation.

Consider now the representation of density $n(r)$ in some basis set. From Bloch theorem, if $\vec{T}$ is the lattice period, we know that n must have the form:

$$\begin{aligned} n_{\vec{k},i}(\vec{r}) &= e^{i\vec{k}\cdot\vec{r}}\phi_{\vec{k},i}(\vec{r}) \\ \phi_{\vec{k},i}(\vec{r}+\vec{T}) &= \phi_{\vec{k},i}(\vec{r}), \quad \vec{k} = \frac{2\pi}{L_x}\vec{i} + \frac{2\pi}{L_y}\vec{j} + \frac{2\pi}{L_z}\vec{k} \end{aligned} \tag{3.36}$$

where i is an energy index. This condition is fulfilled if we choose:

$$\phi_{\vec{k},i}(\vec{r}) = \frac{1}{\sqrt{\Omega}} \sum_{|\vec{G}|\leqslant G_{max}} c_{\vec{k},i} e^{i\vec{G}\vec{r}}, \quad \vec{G} = \frac{2\pi}{L_x}\vec{i} + \frac{2\pi}{L_y}\vec{j} + \frac{2\pi}{L_z}\vec{k} \tag{3.37}$$

where Ω is the box volume. Often only the Γ point $\vec{k} = \vec{0}$ is considered and in other cases efficient sampling of the first Brillouin-zone is used (Monkhorst-Pack grid)[41]. This type of expansion is called a plane-waves expansion and has several advantages:
- the basis doesn't depend on the ions positions;
- only the parameter G_{max} controls the basis size;

The principal disadvantage of this expansion is instead that a large number of plane-waves should be necessary in general. This is due to the fact that plane-waves don't take into account of the atomic structure.

To overcome the problem of large size basis set, pseudopotentials can be used. General ideas which brings to pseudopotentials use are the following:
- to consider the more internal electrons as frozen (frozen-core approximation); - replace the true all electrons potential with a pseudo potential which takes into account the frozen electrons and whose effect on valence electrons is included in the pseudopotential.

Pseudopotentials construction can be a delicate operation and in each case validity of approximations and transferability tests must be done. There are various types of pseudopotentials constructions like H and TM [4, 21, 32, 61] (in the next we will see why these are called norm-conserving pseudopotentials). All norm-conserving pseudopotential constructions have the following properties in common.

Suppose we want to construct a pseudopotential of a N states atom. Of these N states only M are treated as valence electrons while the remaining $N-M$ are frozen. Suppose also to know the all electrons wavefunctions (this in general can be calculated numerically) which satisfy:

$$\begin{aligned}
(-\frac{1}{2}\nabla^2 + v_s)[n]\psi_i(\vec{r}) &= \epsilon_i\psi_i(\vec{r}) \quad i = 1, ..., N, \quad v_s = v_{xc} + v_H + v_{Ie} \\
n_v(\vec{r}) &= \sum_{i=1}^{M} |\psi_i(\vec{r})|^2 \,;\; n_c(\vec{r}) = \sum_{i=1}^{N-M} |\psi_i(\vec{r})|^2 \\
n_{tot} &= n_c + n_v
\end{aligned} \tag{3.38}$$

where n_v, n_c, n_{tot} are respectively the valence, core and total density. In general ψ_i have several nodes. let usindicate with $r_c^i, i = 1, ..., M$ a set of cut-off distances which are near to the last node of $\psi_i, i = 1, ..., M$. We have to generate a set of normalized pseudo wavefunctions ψ_i^{ps} which:
- haven't got nodes;
- are equal to $\psi_i, i = 1, .., M$ for $r^i \geqslant r_c^i$ for some reference energies ϵ_i^{ps} (obviously this condition implies that pseudo and true potential are equal for $r^i \geqslant r_c^i$);
- logarithmic derivatives ($\frac{d}{d|\vec{r}|}\ln\psi$) of all electrons and pseudo wavefunctions are equal at $r^i = r_c^i$.

The fact that pseudo-wavefunctions are normalized and are equal to all-electrons wavefunctions for $r^i \geqslant r_c^i$ implies that (norm conservation):

$$\int\int_0^{r_c^i} r^2|\psi_i(r,\epsilon')|^2 dr d\Omega = \int\int_0^{r_c^i} r^2|\psi_i^{ps}(r,\epsilon')|^2 dr d\Omega =$$

$$= \Big[-\frac{1}{2} r^2 |\psi_i^{ps}(r,\epsilon)|^2 \frac{d}{d\epsilon}\frac{d}{dr} \ln \psi_i^{ps}(r,\epsilon) \Big]_{|r_c^i,\epsilon'} =$$
$$= \Big[-\frac{1}{2} r^2 |\psi_i(r,\epsilon)|^2 \frac{d}{d\epsilon}\frac{d}{dr} \ln \psi_i(r,\epsilon) \Big]_{|r_c^i,\epsilon'} \tag{3.39}$$

Eq. 3.39 means that at r_c^i, not only the logarithmic derivatives but also their first derivative with respect to ϵ are equal. Scattering properties depend on logarithmic derivatives (see [8]), thus this property assures that scattering properties are correct even around the reference energies ϵ_i^{ps}. Inverting the Schrodinger equation in 3.38 an expression for the pseudopotential can be obtained. For example suppose that $i = \{n, l\}$ where l is the angular momenta. Let us write $\psi_{n,l}^{ps}$ as:

$$\psi_{nl}^{ps} = u_{n,l}^{ps}(r) Y_{lm}(\theta, \phi) \tag{3.40}$$

and consequently the pseudopotential (semilocal form) as:

$$\langle \vec{r} | V^{ps} | \vec{r}' \rangle = \sum_n \sum_l v_{n,l}^{ps}(r) |l, m\rangle \langle l', m'| \tag{3.41}$$

where the sum over n, l runs over pseudized states. In polar coordinates the result for the $\{n, l\}$ pseudopotential channel is:

$$v_{n,l}^{ps,scr}(\vec{r}) = \epsilon_{n,l}^{ps} - \frac{l(l+1)}{r^2} + \frac{1}{2u_{n,l}^{ps}} \frac{d^2}{dr^2} u_{n,l}^{ps}(r) \tag{3.42}$$

The superscript "scr" stays screened as it contains the contribution of E_{XC} and E_H of pseudo-valence electrons between themselves. The final unscreened psudopotential is obtained by subtracting the XC and electrostatic contribution:

$$v_i^{ps} = v_i^{ps,scr} - v_H[n_0^{ps}, \vec{r}] - v_{XC}[n_0^{ps}, \vec{r}] \quad n_0^{ps} = \sum_{i=1}^{M} |\psi_i^{ps}(r)|^2 \tag{3.43}$$

then the total energy of the system is:

$$E = K_e[n_v] + V^{ps}[n_v] + E_H[n_v] + E_{XC}[n_v] \tag{3.44}$$

where n is the valence density and E_{XC} refers to the interaction between valence electrons between themselves. The term of E_{XC} between valence electrons and core is included in V^{ps} with a term linear in n_v. E_{XC} in general is

not linear in $n = n_c + n_v$ and if the linear approximation doesn't work non-linear core corrections are needed. This happens if there is a large overlap between n_c and n_v. In these cases a partial-core density, which reproduces the exact core-density for $r \geqslant r_{nlc}$ and is smoother than the exact core density for $r \leqslant r_{nlc}$, is added to the valence density.

What we said until this point is common to all constructions of norm-conserving pseudopotentials. What changes depending on type construction is the ansatz which is done for the pseudowavefunction for $r^{n,l} \leqslant r_c^{n,l}$:
- in hamman construction it is posed $\psi_{n,l}^{ps} = r^{l+1} f_{n,l}(r)$, with:

$$f_{n,l}(r) = c_{n,l}^4 r^4 + c_{n,l}^3 r^3 + c_{n,l}^2 r^2 + c^0 \tag{3.45}$$

Note that the linear term in r is absent such that $\frac{d}{dr}\psi_{n,l}^{ps}|_0 = 0$. These coefficients can be obtained imposing the conditions we said and:

$$\frac{d^m}{dr^m}\psi_{n,l}|_{r_c^{i+}} = \frac{d^m}{dr^m}\psi_{n,l}^{ps}|_{r_c^{n,l-}} \quad m = 1, 2 \tag{3.46}$$

- in Troullier-Martins construction it is posed $\psi_{n,l}^{ps} = r^{l+1} e^{p_{n,l}(r)}$, with:

$$f_{n,l}(r) = c_{n,l}^0 + c_{n,l}^2 r^2 + ... + c_{n,l}^{12} r^{12} \tag{3.47}$$

To determine the extra coefficients used in this construction other extra smoothness conditions are imposed:

$$\begin{aligned} \frac{d}{dr}\psi_{n,l}^{ps}|_0 &= 0 \\ \frac{d^m}{dr^m}\psi_{n,l}|_{r_c^{n,l+}} &= \frac{d^m}{dr^m}\psi_{n,l}^{ps}|_{r_c^{n,l-}} \quad m = 0, ..., 4 \end{aligned} \tag{3.48}$$

Pseudoptentials can be written in more computationally convenient forms in which one separates their local (which are independent of l) from the non local part. The more convenient is the Kleinman-Bylander form. To use KB form one needs some approximations. This can generate "ghost states" (they are not present in the real atom). The presence of such states must be verified when constructing a pseudopotential [23, 24].

It can happen that the all-electrons wavefunction, for some quantum state, doesn't have nodes at $r \neq 0$ and that the only maximum of ψ stays at "small" r (deep-electrons). In this case the pseudization technique here described produces a small decrease of the energy cut-off needed to describe the state. In fact, due to the norm-conservation condition, pseudo and all-electrons wavefunctions are very similar also for $r \leq r_c$. Ultrasoft pseudopotentials were

introduced by Vanderbilt to solve this problem in [35, 67]. In ultrasoft pseudopotentials construction, pseudowavefunctions have norm less than one but satisfy the same properties of norm-conserving pseudopotentials: logarithmic derivatives, matching of eigenvalues, equality to the all electrons wavefunction for $r \geq r_c$. The charge deficit (wavefunctions are not normalized) is made up adding to the valence density an extra-charge term.

3.5 Car-Parrinello molecular dynamics

In section 3.3 we briefly pointed out the various techniques one can use to calculate the electron density $n(\mathbf{r})$ at fixed ion positions. In [9] R. Car and M. Parrinello introduced a method to optimize approximatively wavefuntion on the fly, that is during the atomic motions. With this method the BO surface is not sampled exactly during atomic motions (also optimizing wavefunction at each ionic step one must fix a convergence threshold for orbitals) but orbitals oscillate around the "true" BO surface with an amplitude which has to be bounded during the simulation. CP, introducing a fictitious electronic motion, postulated the following Lagrangian:

$$\begin{aligned}
\mathcal{L}_{CP} &= \mu \sum_i \int |\dot{\psi}_i(\mathbf{r})|^2 d\mathbf{r} + \frac{1}{2}\sum_I M_i \dot{\mathbf{R}}_I^2 - E_{KS}[\psi_i, R_I] + \\
&+ \sum_{ij} \Lambda_{ij}\Big(\langle\psi_i|\psi_j\rangle - \delta_{ij}\Big) = K_f + \mathcal{L} \\
E_{KS}[\psi_i, R_I] &= \sum_i f_i \int \psi_i^*(\mathbf{r})(-\frac{1}{2}\nabla^2)\psi_i(\mathbf{r})d\mathbf{r} + E_H[n] + \\
&+ E_{xc}[n] + V_{eI}(\psi_i, \mathbf{R_I}) + V_{II}(\mathbf{R_I})
\end{aligned} \tag{3.49}$$

where μ is a parameter called fictitious electron mass, E_{KS} is the KS energy, f_i are the orbitals occupation numbers and Λ_{ij} are the Lagrangian multipliers introduced to ensure the orthonormalization condition of KS orbitals. Constraints are always treated with SHAKE-RATTLE algorithms.

The equations of motion then write:

$$\begin{aligned}
\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \mathbf{R_I}} &= \frac{\partial \mathcal{L}}{\partial \mathbf{R_I}} \quad \rightarrow \quad M_I \ddot{\mathbf{R}}_I = -\frac{\partial E_{KS}}{\partial \mathbf{R}_I} \\
\frac{d}{dt}\frac{\delta \mathcal{L}}{\delta \psi_i^*} &= \frac{\delta \mathcal{L}}{\delta \psi_i^*} \quad \rightarrow \quad \mu\ddot{\psi}_i(\mathbf{r},t) = -f_i H_{KS}\psi_i(\mathbf{r},t) + \sum_j \Lambda_{ij}\psi_j(\mathbf{r},t) \\
& \qquad\qquad with \quad \frac{\delta E_{KS}}{\delta \psi_i^*} = f_i H_{KS}
\end{aligned} \tag{3.50}$$

These equations can be integrated with methods of classical molecular dynamics [65, 66] also using norm-conserving or ultrasoft pseudopotentials [36]. The

idea underlying these equations is that classical equations of motion move particles towards energy minimums that is they try to minimize potential energy. Oscillations amplitude around potential energy minimums depend on the fictitious electronic temperature T_f, then if one wants to have small amplitude oscillations, in order to have a good approximation of BO surface, the fictitious electronic temperature has to be "small". In other words let us suppose we apply a thermostat to the ions, then:

$$e^{-\beta H_{CP}} = e^{-\beta(K_f+H)} \sim e^{-\beta H} \qquad if \;\; T_f \ll T \tag{3.51}$$

Moreover let us indicate with ω_I^{max} the maximum ionic frequency and with ω_e^{min} the minimum fictitious electronic frequency. We have (see [50]):

$$\omega_e^{min} \propto \frac{E_{gap}}{\mu} \tag{3.52}$$

where E_{gap} is the smallest energy gap between occupied states or occupied-unoccupied states. If the electrons must be able to follow ions during their motion then must be:

$$\frac{\omega_e^{min}}{\omega_I^{max}} \gg 1 \tag{3.53}$$

From these two observations it follows that the fictitious electronic mass must be "small". Unfortunately the time-step integration length decreases as $1/\mu$ so an ad-hoc value for μ must be chosen.

There are also cases in which E_{gap} can be very small as for example in metal simulations. If E_{gap} becomes small one observes an increase in the fictitious electron kinetic energy. In these cases one can thermostat the electronic subsystem at a temperature much smaller than the ionic system one [7]. However the target electron kinetic must be carefully chosen because a too small kinetic energy should violate adiabadicity of ionic motion with respect to electronic one.

In [59] the error on the BO trajectories was estimated in the rigid ion model. In this model each ion carries its own electronic orbitals which move rigidly with the ion itself. Taking into account the orbitals inertia to atomic motion in [59] was shown that atomic trajectories satisfy:

$$(M_I + \Delta M_I)\ddot{R}_I = F_I^{BO} \qquad \Delta M_I = \frac{2}{3}\mu \sum_i \langle \psi_i| - \frac{1}{2}\nabla^2 |\psi_i\rangle \tag{3.54}$$

where F^{BO} are the BO forces. These equations show that CP equations introduce an effective ion mass. Then if one wants to calculate dynamical quantities

a mass renormalization is needed. Moreover in this approximation if:

$$|| \, |\delta\psi_i\rangle_{CP} || \ll || \, |\psi_i\rangle_{BO} || \quad where \;\; |\psi_i\rangle_{CP} = |\psi_i\rangle_{BO} + |\delta\psi_i\rangle_{CP} \tag{3.55}$$

it can be shown that the fictitious electron kinetic energy writes:

$$K_f = \mu \sum_i \langle \dot{\psi}_i | \dot{\psi}_i \rangle = \frac{1}{2} \sum_I \Delta M_I \langle v_I^2 \rangle = 2 k_B T \frac{\mu}{M_I} \sum_i \langle \psi_i | - \frac{1}{2} \nabla^2 | \psi_i \rangle \tag{3.56}$$

Since in this case electrons move with ions, their motion is sufficiently fast to follow ions. Then this value of K_f can be used as a guideline in the choice of the target fictitious electrons kinetic energy when electrons have to be thermostatted. In [7] a value at least two times larger was suggested.

Thermalization of electronic subsystem can be used to be sure that energy transfer between ionic and electronic subsystem doesn't verify. In these cases one needs to thermalize ionic and electronic system with different temperatures. When different temperatures are used it is not possible to obtain an analytical expression of invariant measure using Eq. 2.5, 2.23. What one would like to sample with the two thermostats is:

$$d\mu(\Gamma) = e^{-\beta_f K_f} \, e^{-\beta H} \tag{3.57}$$

This is true if the interaction between ionic and electronic subsystem is "small". With this approximation the measure can be obtained directly from Eq. 2.5 and has the desired form. In conclusion when thermostatting electronic subsystem the absence of energy transfer is a test that the desired sampling is happening.

3.6 Quantum ion corrections via molecular dynamics

Until now, ions were always considered as classical objects. In some cases quantum corrections to this approximation are important and must be explicitly taken into account. A method to calculate these corrections makes use of the expression of partition function in terms of path integrals whichever way forces are calculated (for example from ab-initio principles). A very comprehensive review on the following results can be found in [63]. For simplicity consider the case of a single atom in an external potential ϕ. it is easy to obtain the following expression for partition function $Z(\beta)$:

$$\begin{aligned} Z(\beta) &= \lim_{P\to\infty} Z_P(\beta) \\ Z_P(\beta) &= \left(\frac{mP}{2\pi\beta\hbar^2}\right)^{P/2} \int dx_1 \cdots dx_P \; e^{-\beta U_{eff}} \\ U_{eff} &= -\sum_{i=1}^{P} \left(\frac{1}{2} m\omega_P^2 (x_{i+1}-x_i)^2 + \frac{1}{P}\phi(x_i)\right)_{x_1=x_P}, \; \omega_P = \frac{\sqrt{P}}{\beta\hbar} \end{aligned} \tag{3.58}$$

Given an observable O depending only on x, from Eq. 3.58 it is easy to obtain:

$$\begin{aligned} \langle O \rangle &= \frac{Tr(O\, e^{-\beta H})}{Tr(e^{-\beta H})} = \lim_{P\to\infty} \left(\frac{mP}{2\pi\beta\hbar^2}\right)^{P/2} \int dx_1 \cdots dx_P \; e^{-\beta U_{eff}} \, O^P_{est}(x_1, ..., x_P) \\ O^P_{est} &= \frac{1}{P} \sum_{i=1}^{P} O(x_i) \end{aligned} \tag{3.59}$$

O^P_{est} is an estimator of the observable O at fixed P. Estimators, which depend only on x can also be constructed for total energy, virial and so on. At fixed P let us introduce P fictitious momenta $p_1, ..., p_P$ and consider the following evolution equations:

$$\begin{aligned} \dot{p}_i &= -\frac{\partial}{\partial x_i} U_{eff}(x_1, ..., x_p) - \frac{p_\eta}{Q} p_i \\ \dot{x}_i &= \frac{p_i}{m} \\ \dot{\eta} &= \frac{p_\eta}{Q} \\ \dot{p}_\eta &= \sum_{i=1}^{P} \frac{p_i^2}{m} - N k_B T \end{aligned} \tag{3.60}$$

These equations represent the time evolution of a close polymer composed of P particles which interact through harmonic potentials in an external field $\frac{1}{P}\phi$, thermalized at temperature T. Given an observable whose estimator depends only on x and under ergodic hypothesis for the system 3.60, for what said in section 2.3 it follows:

$$\left(\frac{mP}{2\pi\beta\hbar^2}\right)^{P/2}\int dx_1\cdots dx_P\; e^{-\beta U_{eff}}\, O^P_{est}(x_1,...,x_P) = \lim_{t\to\infty}\frac{1}{t}\int_0^t O^P_{est}(\tau)d\tau \tag{3.61}$$

where x_i evolve through 3.60. In practice given an observable estimator one has to reach a plateau in time-average value for increasing P. In such form of evolution equations various convergence problem appear increasing P:
1) time evolution of not thermostatted system is of quasi integrable form since the harmonic terms become dominant on the external potential increasing P. The magnitude of the external potential decreases as $1/P$. As was pointed out in section 2.1 it can be shown that these types of systems are not ergodic.
2) Since the harmonic frequencies increase as P an always shorter tim-step length is needed. This problem can be solved recasting the partition function and then time evolution so as to decouple the harmonic forces. In this way, using a multiple time step algorithm, the harmonic part of evolution, which is the faster one, can be analytically calculated. At fixed P increasing the temperature the harmonic term decreases and the problem of quasi integrability reduces. In any case the convergence in P of observable averages can be difficult.

Generalization of this technique to many particle systems can be done. Even for identical particles such a scheme can be used [63] but obviously also this technique is affected by the "sign problem" in the case of fermions.

Instead of starting from the full quantum partition function one can also try to expand the quantum partition function in series of $\hbar$. In some cases one should need only few terms in $\hbar$. In practice one has to check the convergence in function of powers of $\hbar$ for the desired observables. The correction term to classical partition function can be quite easily computed analytically, for example in the Wigner transform formalism [70]. To fix ideas let us consider the h^2 correction. It can be shown (see [2] and references therein) that to order $\hbar^2$ the partition function can be written as:

$$Q_{NVT} \propto \int dr_1\cdots dr_{3N}\Big[1-\frac{\beta^3\hbar^2}{24m}\sum_{i=1}^{3N}\Big((\nabla_{r_i}V(r_1,...,r_{3N})\Big)^2\Big]e^{-\beta V(r_1,...,r_{3N})} \tag{3.62}$$

Higher order terms contain higher powers of β. Increasing the temperature, the higher order terms become less important. Eq. 3.62 can be rewritten as:

$$Q_{NVT} \propto \int d\Gamma\, e^{-\beta V_{eff}}, \quad V_{eff} = V - \frac{1}{\beta} \ln \left[1 - \frac{\beta^3 \hbar^2}{24m} \sum_{i=1}^{3N} \left(\nabla_{r_i} V(r_1, ..., r_{3N}) \right)^2 \right] \tag{3.63}$$

Note that Wigner transform formalism gives us an expression for the power series expansion in $\hbar$ of partition function but doesn't say nothing about power series expansion of dynamics. In fact the Wigner transform of a quantum observable O (function of (r, p)) of a N particles system is defined as:

$$O_W(r, p) = \frac{1}{(2\pi\hbar)^{3N}} \int dz e^{iP \cdot z\hbar} \Big\langle r - \frac{z}{2} \Big| \widehat{O} \Big| r + \frac{z}{2} \Big\rangle \tag{3.64}$$

Time evolution can be written as:

$$i\hbar \frac{\partial}{\partial t} O_W(r, p) = \{O_W, H_W\}_M = O_W * H_W - H_W * O_W \tag{3.65}$$

where $\{\}_M$ stays for Moyal brackets and:

$$(\widehat{A}\widehat{B})_W = A_W * B_W = A_W \exp\left(i \frac{\hbar}{2} \Lambda\right) B_W \tag{3.66}$$

$$\Lambda = (\overleftarrow{\nabla}_r \overrightarrow{\nabla}_p - \overleftarrow{\nabla}_p \overrightarrow{\nabla}_r) \tag{3.67}$$

For r and p one simply obtains:

$$\dot{r}_W = \frac{p_W}{m}$$
$$\dot{p}_W = -\frac{\partial}{\partial r} V_W(r) \tag{3.68}$$

Since Wigner's transform of an observable function of r only is simply the function evaluated at r, the last equation is saying nothing about power series $\hbar$ expansion of dynamics. Conversely a power series in $\hbar$ for the equilibrium canonical partition function can be obtained simply calculating to desired order:

$$[e^{-\beta \widehat{H}}]_W = \sum_{k=0} \frac{-\beta^k}{k!} [\widehat{H}^k]_W \tag{3.69}$$

using 3.66. We can however sample this measure via fictitious MD equations. The following Hamiltonian must be used to generate MD trajectories:

$$\begin{aligned} H_{eff} &= T + V_{eff}, \quad V_{eff} = V - \frac{1}{\beta}\ln\left[1 - \frac{\beta^3\hbar^2}{24m}\sum_{i=1}^{3N}\left(\nabla_{r_i}V(r_1,...,r_{3N})\right)^2\right] \\ \dot{q}_i &= \frac{p_i}{m} \\ \dot{p}_i &= -\frac{\partial V_{eff}}{\partial r_i} \end{aligned} \tag{3.70}$$

In fact coupling this system to a NH thermostat the following equations are obtained:

$$\begin{aligned} \dot{q}_i &= \frac{p_i}{m} \\ \dot{p}_i &= -\frac{\partial V_{eff}}{\partial r_i} - \frac{p_i p_\eta}{Q}, \\ \dot{\eta} &= \frac{p_\eta}{Q} \\ \dot{p}_\eta &= \sum_{i=1}^{3N}\frac{p_i^2}{m_i} - dNk_BT \end{aligned} \tag{3.71}$$

In section 2.3 it was shown that these equations sample the following measure:

$$d\mu(\Gamma) = d\Gamma\, e^{-\beta H_{eff}} \tag{3.72}$$

Then modifying equations of motion in this way, all $\hbar$ correction terms can be included. Moreover not ergodicity problems are present in this case. To test this method I have calculated the specific heat of a one dimensional harmonic oscillator. In this case, taking $\omega = m^{-\frac{1}{2}}$, H_{eff} simply writes:

$$H_{eff} = \frac{p^2}{2m} + \frac{1}{2}q^2 - k_BT\ln[1-\alpha_2 q^2], \quad \alpha_2 = (24mk_B^3T^3)^{-1} \tag{3.73}$$

coupling this system to a NH chain thermostat, the following equations are obtained:

$$\begin{aligned} \dot{q} &= \frac{p}{m}, \quad \dot{\eta}_i = \frac{p_{\eta_i}}{Q} \quad i = 1,...,M \\ \dot{p} &= -q - \frac{\alpha_1 q}{1-\alpha_2 q^2} - \frac{p\, p_{\eta_1}}{Q_1}, \quad \alpha_1 = (12mk_B^2T^2)^{-1} \end{aligned}$$

$$\dot{p}_{\eta_1} = \frac{p^2}{m} - k_B T - \frac{p_{\eta_2} p_{\eta_1}}{Q_2}, \quad \dot{p}_{\eta_k} = \frac{p^2_{\eta_{k-1}}}{Q_{k-1}} - k_B T - \frac{p_{\eta_{k+1}}}{Q_{k+1}} p_{\eta_k} \quad k = 2, ..., M-1$$
$$\dot{p}_{\eta_M} = \frac{p^2_{\eta_{M-1}}}{Q_{M-1}} - k_B T \tag{3.74}$$

The average total energy of the system has to be calculated through the formula:

$$\langle E \rangle = -\frac{\partial}{\partial \beta} \ln Q = -\frac{\partial}{\partial \beta} \ln \Big[\int e^{-\beta H_{eff}(q,p,\beta)} dq dp \Big] =$$
$$= \frac{1}{Q} \int \Big[H_{eff} + \beta \frac{\partial}{\partial \beta} H_{eff} \Big] e^{-\beta H_{eff}} dq dp =$$
$$= \frac{1}{Q} \int \Big[\frac{p^2}{2m} + \frac{q^2}{2} + \frac{3}{2} \Big(\frac{\alpha_1 q^2}{1 - \alpha_2 q^2} \Big) \Big] e^{-\beta H_{eff}} dq dp \tag{3.75}$$

This means that the function:

$$Est[E] = \frac{p^2}{2m} + \frac{q^2}{2} + \frac{3}{2} \Big(\frac{\alpha_1 q^2}{1 - \alpha_2 q^2} \Big) \tag{3.76}$$

is the estimator of total energy and has to be time averaged with (q, p) evolving through 3.74. Using thermodynamics relations the estimators of any observable can be obtained. Note that the estimator of $\langle E \rangle$ does not depend on logarithm of some quantity, which is a good property. Moreover exactly the same procedure can be used in presence of corrections depending on p with an appropriate redefinition of dynamics and estimators.

For high T,since $q \simeq \sqrt{k_b T}$, $Est[E] \simeq \frac{p^2}{2m} + \frac{q^2}{2} = H$, then the following formula can be used to estimate the heat capacity C_V:

$$C_V = \frac{\langle H^2 \rangle_{eff} - \langle H \rangle^2_{eff}}{k_B T^2} \tag{3.77}$$

where:

$$\langle O \rangle_{eff} = \lim_{t \to \infty} \frac{1}{t} \int_0^\infty O(q(\tau), p(\tau)) d\tau \tag{3.78}$$

and $(q, p)(\tau)$ is a solution of Eq. 3.74.

The system parameters were chosen as follows:
$M = 2$, $m = 1836$ $a.u.$ (Hydrogen mass), $Q_1 = Q_2 = 50$ $a.u.$, $dt = 0.08\sqrt{\frac{Q}{2k_B T}}$, $\eta_1(0) = \eta_2(0) = 0$, $p_{\eta_2}(0) = 0$, $p_{\eta_1}(0) = \sqrt{2k_B T Q}$, $q(0) = \sqrt{k_B T}$, $p(0) = \sqrt{m k_B T}$. The integration algorithm was derived using Liouville formalism and can be found in [20]. Obtained results are reported in Fig. 3.1.

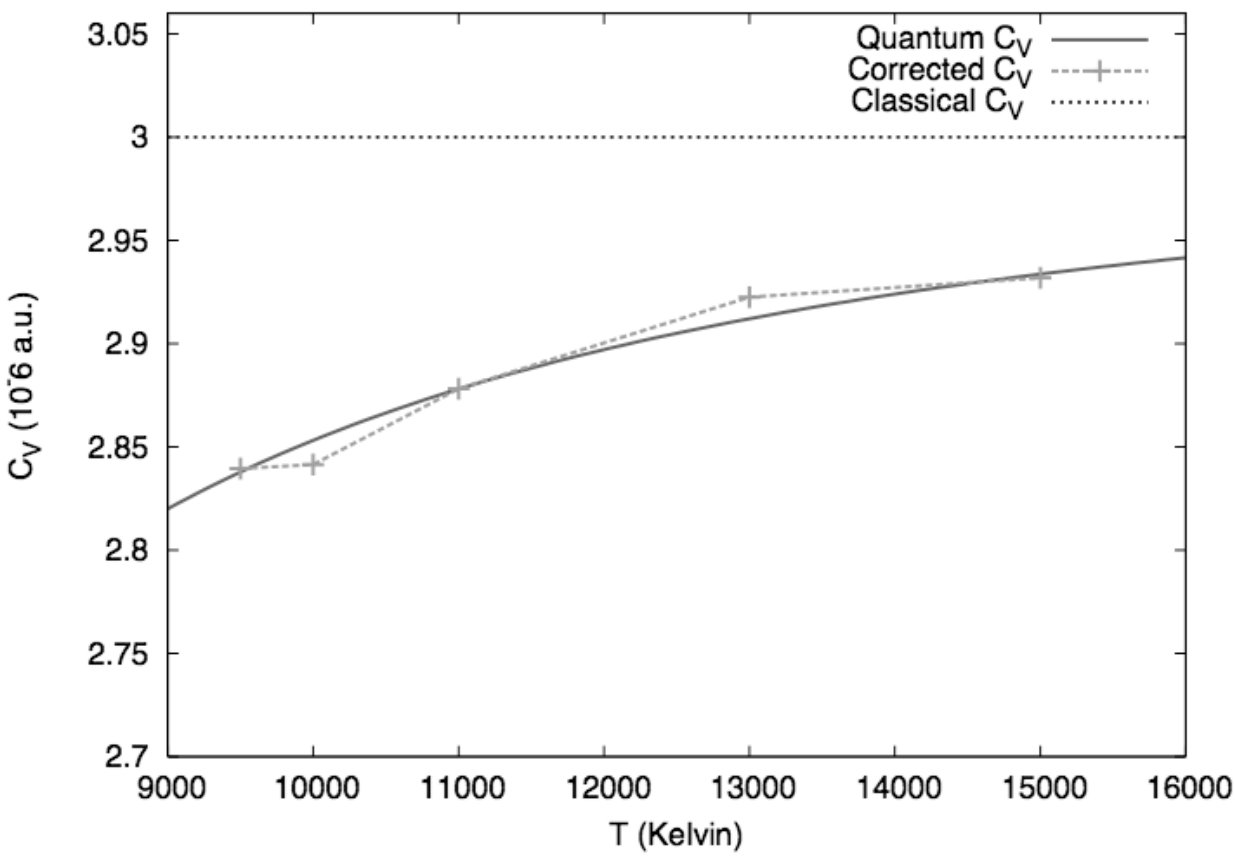


Figure 3.1: Heat capacity C_V as function of temperature for harmonic oscillator. Quantum is the value calculated using quantum statistical mechanics, classical is the classical statistical mechanics result and corrected is the value obtained including the $O(\hbar^2)$ quantum corrections ($k_B \simeq 3 * 10^{-6}$ $a.u.$).

4 Calculations with norm conserving pseudopotentials

In this section we present the first ab-initio calculations of Chromium and Carbon. All calculations reported in this section were done with "AbInit" code and the pseudopotential generator code was "fhi98pp". The program generates norm-conserving pseudopotentials and contains a library of previous constructed pseudopotentials. At first, some tests for both C and Cr pseudopotentials were done. C library pseudopotential tests seemed to be quite correct. Conversely we decided to rebuild Cr pseudopotential since some physical properties obtained with library pseudopotential were poorly reproduced (see next). All total energy calculation reported in this section were done using the BFGS geometry optimization algorithm which is a quasi-Newton method different from the conjugate gradient minimization algorithm (a review of various optimization techniques can be found in [52]). All calculations, except those differently specified, were done without spin polarization corrections.

In all optimized geometry of lattice structures (perfect graphite, Cr crystal, ...), where *acell* (the length of primitive cell in the three spatial directions) is optimized and atom positions are fixed, the convergence criteria chosen for forces is $5d^{-5}$ Ha/Bohr. For geometries different from these, like those reported in section 4.6, 4.7 the threshold is $5d^{-4}$ Ha/Bohr.

4.1 Carbon pseudopotential test

The carbon library pseudopotential is of the TM type with $2s^2$ and $2p^2$ as valence states. At first we tried to obtain the following experimental graphite data:

$$a = b = 2.46 \text{ Å } c = 3.41 \text{ Å}, \quad E_{ch} = -7.4 \text{ eV} \tag{4.1}$$

where the cohesive energy E_{ch} of a crystal structure whose primitive cell has N atoms is defined as:

$$E_{ch} = E_{tot}/N - E_{atom} \tag{4.2}$$

All energies are calculated at $T = 0$ and E_{atom} is the energy of an isolated atom of the same substance.

We confronted results with respect to the type of XC functional used. The use of a GGA functional produces some improvements of results and this type of functional will be used in all other calculations. In particular we used the PBE functional of GGA type and compare it with PW92 functional of LDA type. The value of cohesive energy is better estimated if spin-polarized calculations are done. To test the quality of carbon pseudopotential, the lattice parameters and cohesive energy of bulk graphite have been calculated after a convergence study with respect to k-points grid and energy cut-off. The primitive cell of graphite has four atoms and is of hexagonal type (Table 4.1).

Graphite primitive cell			
acell(Bohr)	4.649	4.649	12.627
rprim	1.0	0.0	0.0
	-0.5	$\sqrt{3}/2$	0.0
	0.0	0.0	1.0
xred	0.0	0.0	0.0
	2/3	1/3	0.0
	0.0	0.0	0.5
	1/3	2/3	0.5

Table 4.1: Input variables for the primitive cell of graphite: *acell* corresponds to the length of primitive cell in the three spatial directions; *rprim* are the director vectors of the primitive cell and *xred* the scaled positions of atoms.

In order to estimate the cohesive energy, the total energy of an isolated carbon atom has to be calculated. Convergence study with respect to the box

dimension was done (Tab. 4.2). The convergence study with respect to the

acell(x,y,z) (Bohr)	Total energy (Ha)
16.00	-5.35907
18.00	-5.35900
20.00	-5.35900
22.00	-5.35900

Table 4.2: Results of convergence study with respect to box dimension for an isolated carbon atom, cut-off 40 Ha.

energy cut-off for graphite lattice was done at first using a grid of 4 4 4 k-points. As can be seen in Tab. 4.3, the changes of calculated lattice parameters and total energy are very small even if with a cut-off of 30 Ha. Fixing the cut-off at 40 Ha the convergence with respect to k-points grid was studied (Tab. 4.4). Also using the 4 4 2 grid the total energy error is of abut 19 meV with respect to the 6 6 6 grid. Moreover the lattice parameters error with respect to its experimental value is of 0.3%.

Cut-off (Ha)	a, b, c (Bohr)			E_{ch} (Ha)	ΔE_{ch} (eV)
30	4.6476	4.6476	12.6230	-0.34598	0.000
35	4.6474	4.6474	12.6223	-0.34613	-0.0041
40	4.6471	4.6471	12.6215	-0.34627	-0.0038
45	4.6473	4.6473	12.6222	-0.34638	-0.0030

Table 4.3: Results of convergence study of graphite with respect to cut-off energy .

k-points	a, b, c (Bohr)			E_{ch} (Ha)	ΔE_{ch} (eV)
4 4 2	4.6474	4.6474	12.6224	-0.34529	0.0000
4 4 4	4.6474	4.6474	12.6223	-0.34599	-0.0190
5 5 5	4.6508	4.6508	12.6316	-0.34509	0.0245
6 6 6	4.6511	4.6511	12.6325	-0.34470	0.0106

Table 4.4: Results of k-points convergence study of graphite with 40 Ha energy cut-off .

The final estimated values of cohesive energy and lattice parameters are:

$$a = b = 2.458 \text{ Å}, \; c = 3.35 \text{ Å}, \;\; E_{ch} = -9.2 \text{ eV} \tag{4.3}$$

Using an LDA XC functional and pseudopotential was obtained:

$$a = b = 2.448 \text{ Å}, \; c = 3.30 \text{ Å} \;\; E_{ch} = -10 \text{ eV} \tag{4.4}$$

Then we chose to use GGA corrections in our calculations.

Also the band structure of graphite has been calculated (Fig. 4.1). In general, both LDA and GGA approximations tend to overestimate the cohesive energy of crystals. Also in our case, while the lattice parameters error is very small, the cohesive energy error is quite large (about 25%).

In this section we referred to spin restricted calculations. This approximation is correct for graphite which isn't a magnetic material but isn't correct for isolated carbon atom which has the two $2p$ electrons with the same spin. The spin unrestricted calculation for the carbon atom reduces the cohesive energy error. The obtained values are:

$$E_{ch} = -8.8 \text{ eV } in\ LDA \qquad E_{ch} = -7.9 \text{ eV } in\ GGA. \tag{4.5}$$

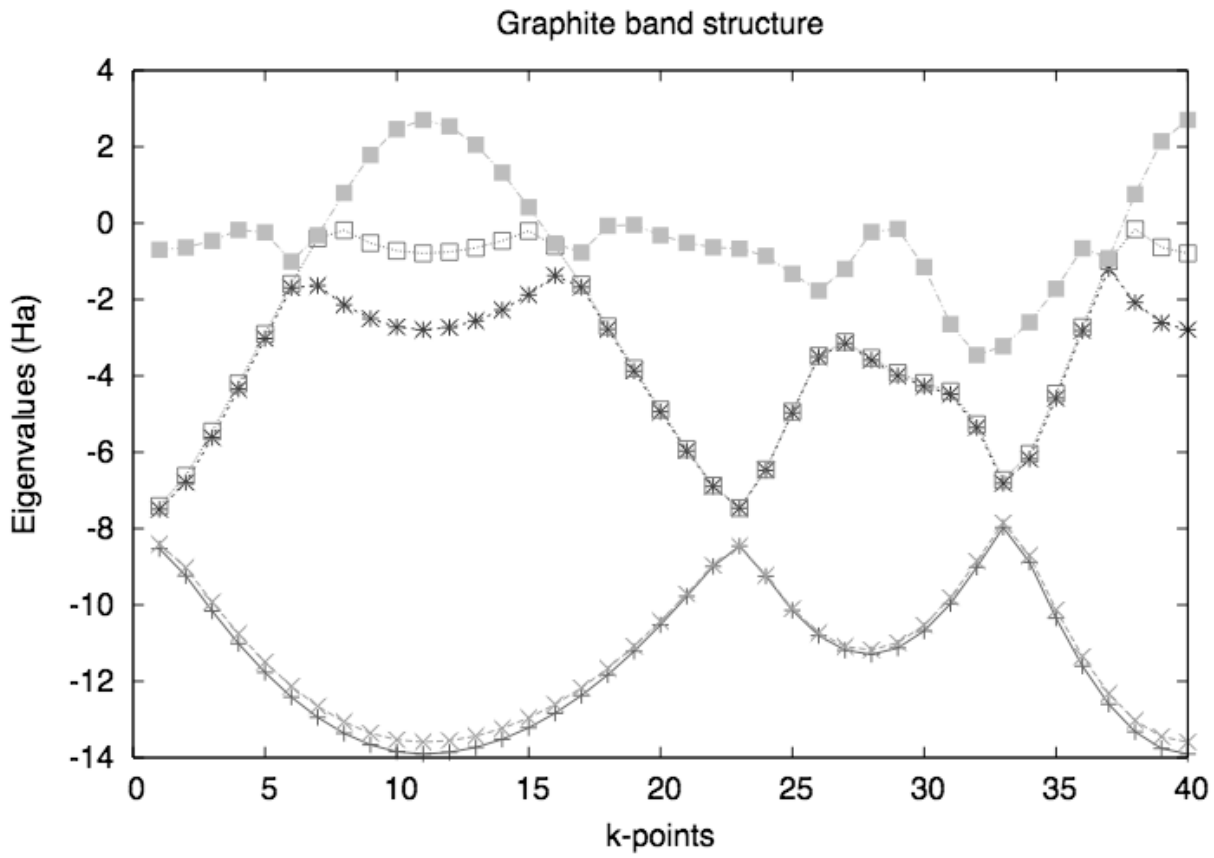


Figure 4.1: Band structure of graphite: points 1 to 10 correspond to the first five eigenvalues from the k-point L=(1/2,0,0) to Γ=(0,0,0), points 11 to 22 from Γ=(0,0,0) to X=(0,1/2,1/2) and points 23 to 40 from X to Γ=(1,1,1).

4.2 Graphite four-fold defect

In [60] the formation energy was calculated and bond length of the four-fold spiro-interstitial defect in graphite (Fig. 4.2). If the graphite cell to which a defective atom is added has N atoms, the formation energy is defined as:

$$E_f = E_{def} - E_{N+1} \tag{4.6}$$

We have calculated the same properties of this type of defect and compared them with [60]. The energy cut-off is the same as the previous calculations and only the k-points grid has been changed (Tab. 4.5) since cell dimension and atomic geometries are changed. In this calculation we use a 2x2x1 cell. We chose a 2 2 2 grid and the total energy error with respect to the 4 4 4 grid is of about 0.2%.

At first, keeping all carbon graphite atoms fixed, the total energy of the system was calculated moving the interstitial carbon between the planes (Fig. 4.3-4.4). Then a geometry optimization, starting from the estimated minimum energy position for the interstitial carbon,was done. The bond length and the

k-points	a, b, c (Bohr)			Total Energy (Ha)
2 2 2	9.2942	9.2942	12.6215	-91.28434
3 3 3	9.3015	9.3015	12.6314	-91.26559
4 4 4	9.3000	9.3000	12.6294	-91.27560

Table 4.5: Convergence study with respect to k-points for a 2x2x1 graphite cell.

formation energy calculated are:

$$E_f = 7.2 \text{ eV}, \quad d_{bond} = 1.48 \text{ Å} \tag{4.7}$$

These values are in accordance with those obtained in [60] with a 4x4x1 cell:

$$E_f = 7.0 \text{ eV}, \quad d_{bond} = 1.5 \text{ Å} \tag{4.8}$$

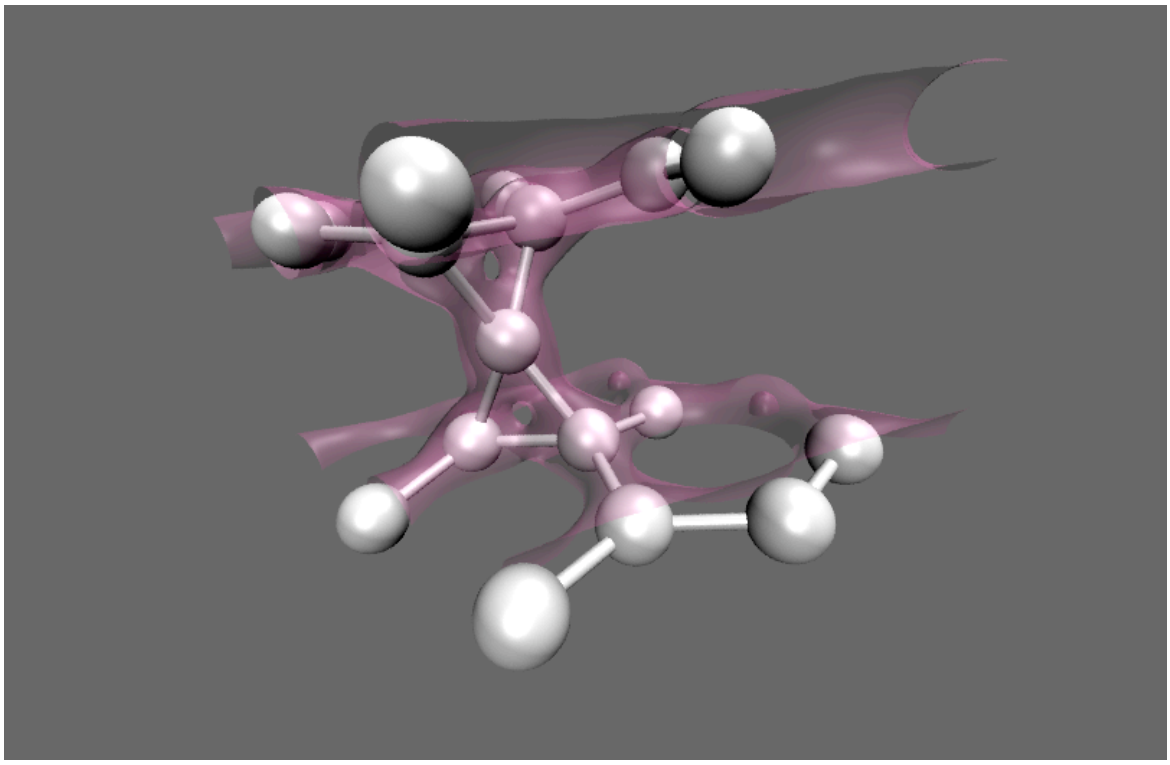


Figure 4.2: Isodensity surface ($\rho(r) \simeq 0.2$) for a 2x2x1 graphite cell with the interstitial four-fold coordinated carbon defect.

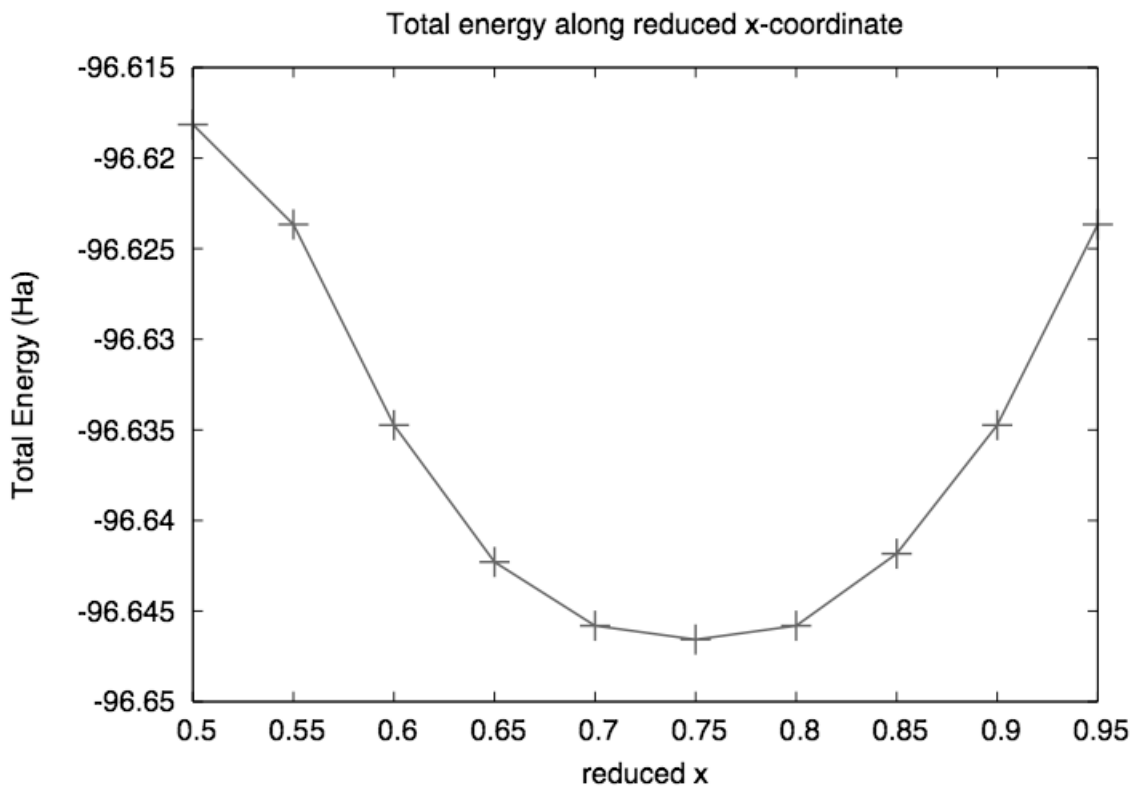


Figure 4.3: Total energy of the system vs reduced x coordinate of the interstitial carbon atom.

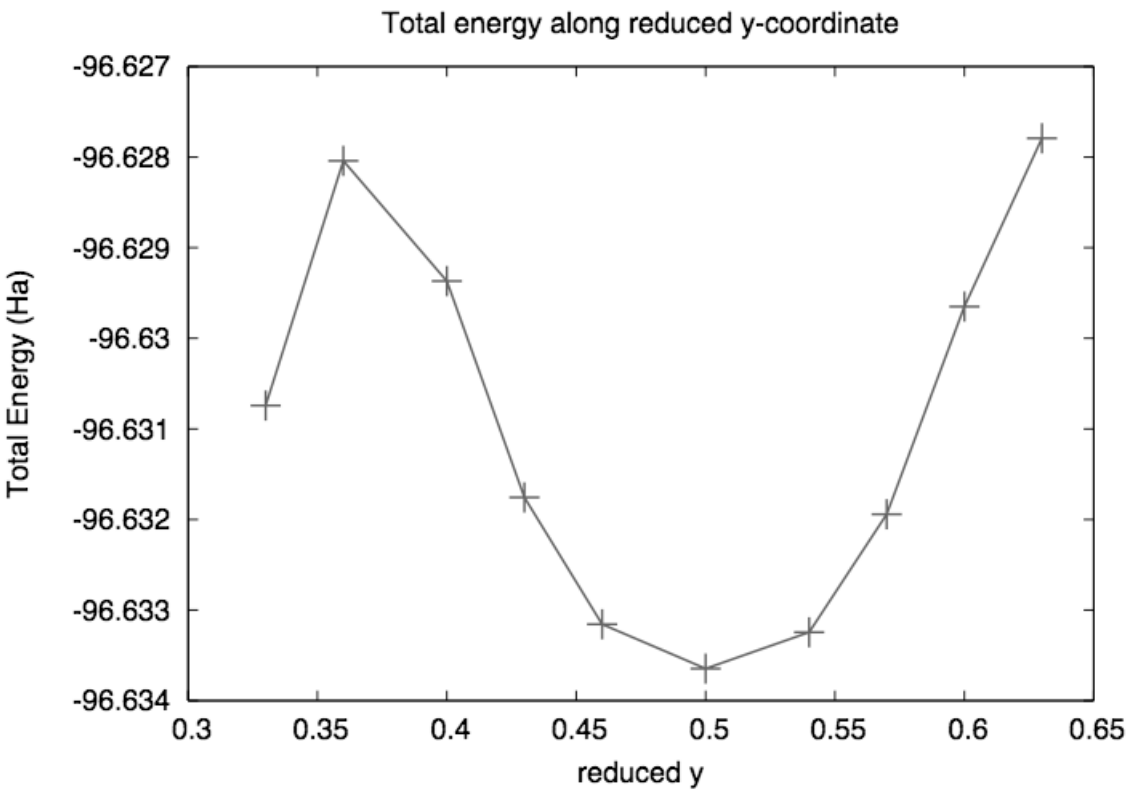


Figure 4.4: Total energy of the system vs reduced y coordinate of the interstitial carbon atom.

4.3 Construction of Chromium pseudopotential

For Chromium some tests of a library pseudopotential lead us to rebuild a new pseudopotential. In this pseudopotential NLCC are not present and semicore states are all included in the core.

After convergence studies with respect to k-points grid and energy cut-off (50 Ha) the lattice parameters and cohesive energy of bulk Chromium were calculated. The lattice parameters are very accurate:

$$a, b, c = 5.46 \text{ Bohr}, \quad 5.44 \text{ Bohr Experimental} \tag{4.9}$$

The cohesive energy instead is grossly overestimated. In a spin restricted calculation we obtained:

$$E_{ch} \simeq -9 \text{ eV}, \quad -4.1 \text{ eV Experimental} \tag{4.10}$$

If an unrestricted calculation is done E_{ch} results positive. In fact even if the spin correction has no effect in the bulk calculus, as reported in Tab. 4.6 (we've checked different initial spin states), the total energy of an isolated Cr atom is of -9.55 Ha while the energy/atom of the crystal is of about -9.51 Ha.

Spin Pol.	a, b, c (Bohr)	Total Energy (Ha)
0	5.4703	-9.51168
1	5.4835	-9.51108
2	5.4703	-9.51168
3	5.4834	-9.51108
4	5.4584	-9.50829
5	5.4827	-9.51108

Table 4.6: Lattice parameters and total energy for different initial spin polarization.

For these reasons we've tried to construct a new pseudopotential.

The used pseudopotentials generation code is "fhi98pp" which can construct TM pseudopotentials in the fully not local KB form. The code can also do a ghost states analysis, logarithmic derivatives calculations and energy cut-off convergence study.

The fact that the total energy of an isolated spin polarized atom is lower with respect to the unpolarized case is general. The difference between these

energies is called spin polarization energy and depends on the pseudopotential used. The spin polarization energy, calculated using pseudopotentials, both in GGA or LDA, is in general bigger with respect to the all electron case.

We cannot calculate the spin polarization energy because there are convergence problem with an AE calculation using "fhi98pp" and PBE functional.

Using the PW91 functional, the spin polarization energy in AE case is of 0.188 Ha. For the library pseudopotential this energy was 0.27 Ha.

The electronic structure of the Cr atom is [Ar]$4s^1 3d^5$ and the $3d$ are the first occupied d orbitals then very deep. In our pseudopotential are not present semicore states. The maximum angular momentum channel is $l = 2$ and the local part is with $l = 0$. Other choices of local part produce ghost states.

Choosing r_{match} =2.56 Bohr for both $l = 1, 2$ channels, a smaller r_{match} for d orbitals (r_d =1.60 Bohr) and taking into account NLCC ($r_{nlc} = 0.8$ Bohr), a better estimation of the spin polarization energy is obtained:

$$E_{spinpol} = 0.208 \text{ Ha} \tag{4.11}$$

Moreover we can do the following considerations:
1) we can decrease the values of (r_d, r_{nlc}) to better estimate $E_{spinpol}$ with the pseudopotential. If we set (r_d, r_{nlc}) to 0.2 Bohr, the spin polarization energy is 0.199 Ha.
2) In general even if these parameters (r_d, r_{nlc}) are very small, the spin polarization energy is still not correct. In the case of Mn, the error that the pseudopotential with very small (r_d, r_{nlc}) with respect to the AE case is of about 5%. The total error is of about 10%.

Correcting our value of the 10% we find 0.1872 Ha which is very close to the LDA AE calculation. The estimated total energy of the spin polarized Cr atom is:

$$E_{Cr} = -17.246455 \text{ Ha} \tag{4.12}$$

In Tab. 4.7 a transferability test is reported for the pseudopotential and in Fig. 4.3 its logarithmic derivatives. In the transferability test we compare eigenvalues energies of pseudopotential with respect to AE calculation in electronic configurations different from the ground state $4s^1 3d^5$. Also fractional electron occupation numbers can be compared.

AE Atom (eV)			Pseudo Atom (eV)	
ConFig.	E_d	E_s	E_d	E_s
$4s^0 3d^5$	-9.8747	-9.9518	-9.8273	-9.9951
$4s^{1/4} 3d^5$	-7.9957	-8.3801	-7.9612	-8.4100
$4s^{1/2} 3d^5$	-6.2056	-6.8635	-6.1836	-6.8817
$4s^{3/4} 3d^5$	-4.5090	-5.4042	-4.4985	-5.4125
$4s^1 3d^5$	-2.9133	-4.0074	-2.9134	-4.0074

Table 4.7: Total energies and first eigenvalues energies of the pseudoatom and AE atom in different configurations.

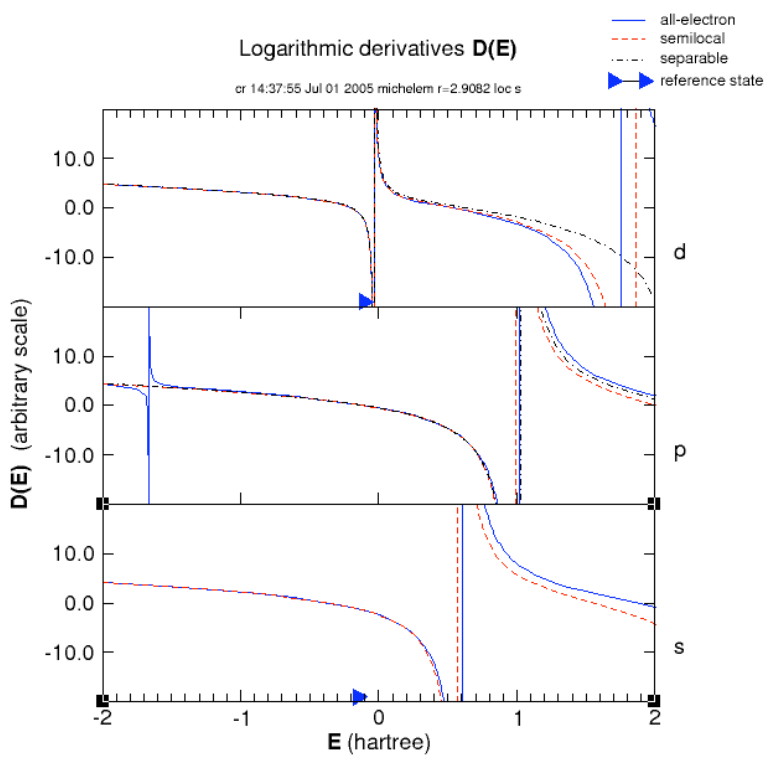


Figure 4.5: Logarithmic derivatives of the AE and pseudoatom in semilocal and fully non-local form.

4.4 Chromium crystal

Chromium has a bcc lattice and only one atom at origin in its primitive cell. Starting with a 8 8 8 grid, we check the convergence with respect to the energy cut-off. The code used to generate the pseudopotential makes a convergence study of the kinetic energy in the k-space. In particular the code calculates the kinetic energy difference, for a given angular momentum, at fixed cut-off with respect to continuum plane waves basis set. The kinetic energy part is often the dominant one but, since NLCC are present, we check again the convergence (Tab. 4.8). The d part is the hardest and an error of 0.1% ($\simeq$ 50 meV) is done with a 50 Ha cut-off for d eigenvalue.

We checked also that 50 Ha is sufficient for the values of lattice parameters which we used to calculate the bulk modulus (4.8-6.0 Bohr) (Tab. 4.9).

Cut-off (Ha)	a, b, c (Bohr)	E_{ch} (Ha)	ΔE_{ch} (eV)
45	5.4380	-0.15170	0.0000
50	5.4328	-0.15309	-0.0378
55	5.4516	-0.15319	-0.0027

Table 4.8: Convergence with respect to the cut-off energy.

Cut-off (Ha)	Total Energy (Ha)
50	-17.34000
55	-17.34011
60	-17.34011

Table 4.9: Total energies for a=b=c=4.8 Bohr in function of the energy cut-off.

With a 50 Ha cut-off we checked the convergence with respect to k-points grid (Tab. 4.10). Dealing with a metallic system the k-point convergence could be very slow. In fact for metals the band structure should not be a sharp function and a very dense grid should be necessary. In this case one can accelerate the convergence using a smearing function for the occupation numbers. The Fermi-Dirac smearing function is not particularly useful to accelerate k-point convergence. Other type of smearing functions were introduced which are

more efficient [68]. Smearing functions, different from Fermi-Dirac function, are characterized by a fictitious temperature parameter. (Tab. 4.11).

k-points	a, b, c (Bohr)	E_{ch} (Ha)	ΔE_{ch} (eV)
10 10 10	5.4540	-0.15328	0.0000
12 12 12	5.4450	-0.15303	0.0070
14 14 14	5.4325	-0.15326	-0.0063
16 16 16	5.4490	-0.15299	0.0072

Table 4.10: Convergence study with respect to the k-points grid.

k-points	a, b, c (Bohr)	E_{ch} (Ha)	ΔE_{ch} (eV)
4 4 4	5.4192	-0.15860	0.0000
6 6 6	5.4475	-0.15193	0.1814
8 8 8	5.4342	-0.15309	-0.0314
10 10 10	5.4527	-0.15316	-0.0019

Table 4.11: Convergence study with respect to the k-points grid with fictitious electronic temperature fixed at 1.0d-2 Ha. The occupation scheme is of Marzari type.

The estimated cohesive energy is:

$$E_{ch} = -4.17 \text{ eV}, \ \simeq -4.1 \text{ eV Experimental} \tag{4.13}$$

The final value of pressure in the cell is of about 4.0d-2 GPa.

The bulk modulus was estimated calculating the total energy in function of the lattice parameter (see Fig. 4.6) through:

$$B = V \frac{\partial^2 E(V)}{\partial^2 V} \tag{4.14}$$

calculated at the minimum of $E(V)$. For a *bcc* lattice $a = b = c$ and $V = \frac{1}{2}a$ then:

$$B_{Cr} = \frac{1}{9a} \frac{\partial^2 E(a)}{\partial^2 a} \tag{4.15}$$

We obtained:

$$B_{Cr} = 244 \text{ GPa}, \ \simeq 170 \text{ GPa Experimental} \tag{4.16}$$

However experimental values refer to a policrystalline situation. Moreover in [47] the chromium bulk modulus was calculated in the LDA approximation and the obtained value is of 250 GPa.

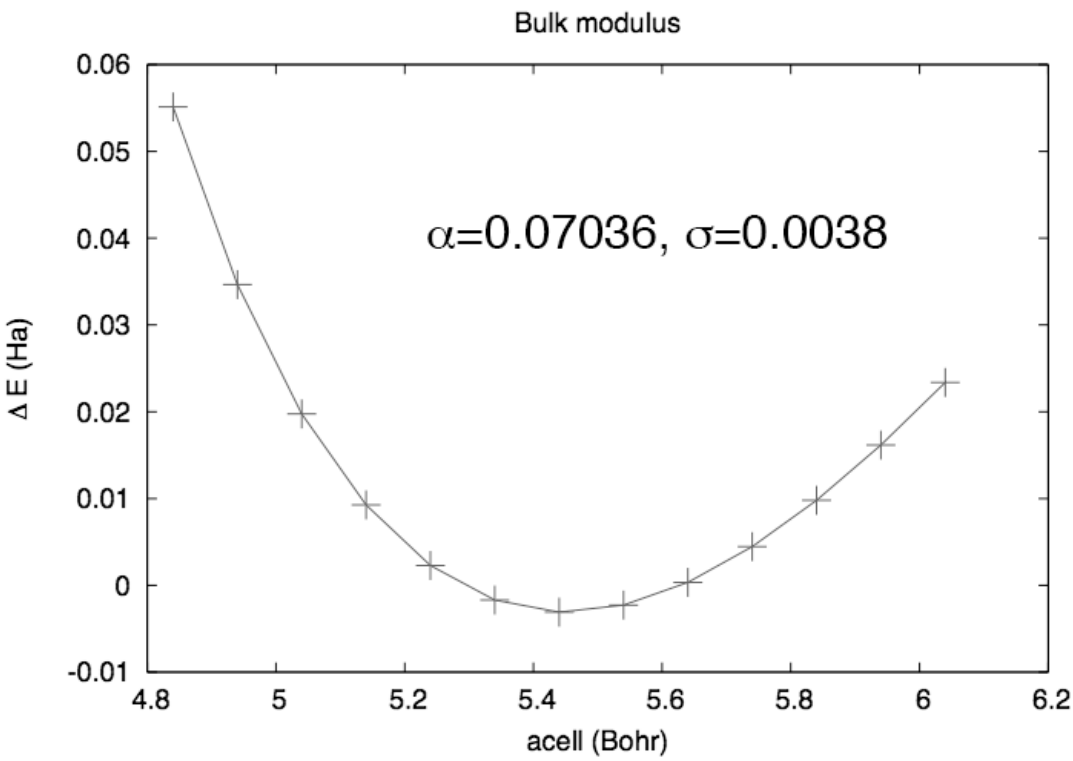


Figure 4.6: Energy difference with respect to *acell* 5.45 vs *acell*, α is the second derivative of ΔE with respect to *acell* (Ha/Bohr^2) and σ its standard deviation.

4.5 Cr-C compounds

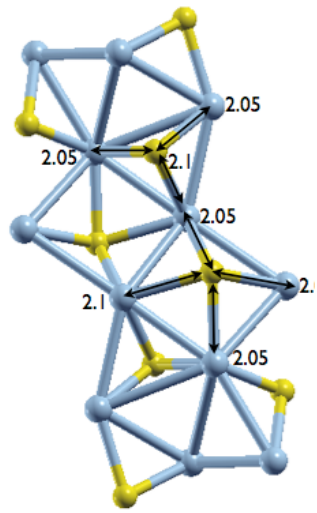


Figure 4.7: Unit cell of Cr_3C_2 carbide, distances are in Å.

To check that the pseudopotential describes correctly the Cr-C interaction we have calculated the lattice parameters of two compounds of Cr and C. From now on the cut-off is fixed at 50 Ha. The simplest carbide is Cr_3C_2, its spacial group symmetry is Pnma 62 and has twenty atoms in the unit cell with:

$$a = 10.4574, b = 5.3468, c = 21.6821\,\text{Bohr} \tag{4.17}$$

The k-points grid is 4 4 4 which, taking into account the symmetries, produces an uniformous k-points density in the three directions.

The obtained lattice parameters are:

$$a = 10.5072, b = 5.3723, c = 21.6853\,\text{Bohr} \tag{4.18}$$

The error is of 0.5%. We also calculated the lattice parameters of CrC with a NaCl type structure [56]. The calculated lattice parameters are:

$$a = b = c = 7.8216\,\text{Bohr}, \quad 7.6167\ \text{Bohr Experimental} \tag{4.19}$$

and in this case the error is of about 2.5%. In [56] the calculated value is 7.5789 Bohr and the LDA approximation was used. However in some cases

the LDA approximation reproduces better than GGA the lattice parameters. In fact using an LDA pseudopotential we obtained

$$a = b = c = 7.6223 \text{ Bohr} \tag{4.20}$$

Both in the CrC and in Cr_3C_2 (Fig. 4.7) the distances between carbon and chromium atoms are of about 2 Å.

4.6 Graphene surface and chromium

In this section are reported the results of the adhesive properties of a Cr atom on a perfect graphene surface. The cell dimension is like a 2x2x1 graphite cell but only the plane at $z = 0$ is present and k-points grid is 2x2x2.

Five trial configurations were chosen (Fig. 4.8-4.9). At first the z positions of Cr which minimize the total energy were simply estimated moving the Cr atom along the z axis. The obtained plane-Cr distances are all around 3.4 Bohr (Fig 4.10). Subsequently the geometry was optimized in the five configurations. The results can be summarized as follows:
- configurations 4 and 5 move towards the configuration 3;
- in the configurations 1, 2, 3 the Cr atom binds strongly to the surface.
Since starting from configurations 3, 4 or 5 the same final configuration is reached, only the configuration 4 was carried to convergence (see Tab. 4.12) Reference energy is the sum of energy of an isolated Cr atom plus the energy of a graphene surface in a 2x2x1 cell.

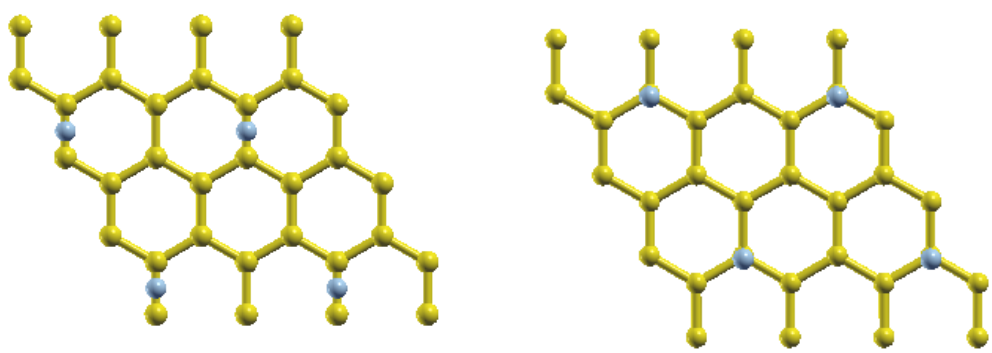


Figure 4.8: On the left the configuration 1; on the right the configuration 2. In all cases the initial Cr-plane distance is 3.4 Bohr.

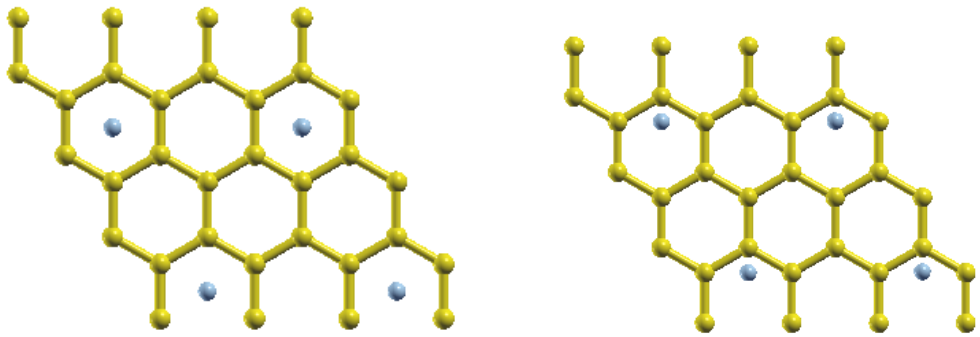
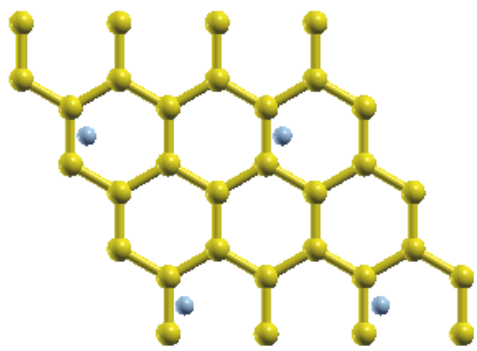


Figure 4.9: Above, on the left the configuration 3; on the right the configuration 4. Below, configuration 5. In all cases the initial Cr-plane distance is 3.4 Bohr.



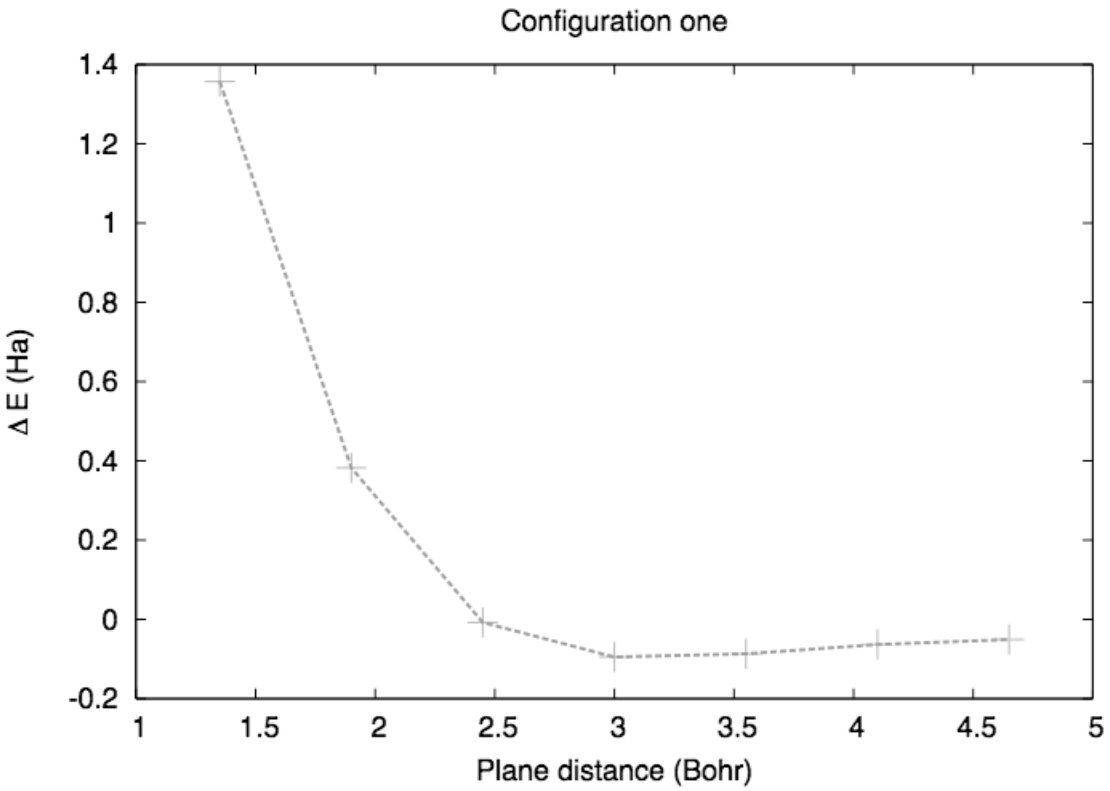


Figure 4.10: Total energy differences for different plane-Cr distances in the configuration 1. For the other configurations similar curves were obtained.

Conf.	Final Distance (Bohr)	Final Stress(Ha/Bohr^3)		ΔE (Ha)	Pressure (GPa)
1	3.8046	σ_{11}= -1.37E-04	σ_{32}= 0.0E+00	-0.12731	2.0
		σ_{22}= -7.7E-05	σ_{31}= 0.0E+00		
		σ_{33}= 9.E-06	σ_{21}= 0.0E+00		
2	3.7186	σ_{11}= -8.9E-05	σ_{32}= 1.E-06	-0.13556	1.7
		σ_{22}= -9.3E-05	σ_{31}= 0.0E+00		
		σ_{33}= 1.E-05	σ_{21}= 0.0E+00		
3	3.0940	σ_{11}= -1.73E-04	σ_{32}= 0.0E+00	-0.14203	3.1
		σ_{22}= -1.59E-04	σ_{31}= 0.0E+00		
		σ_{33}= 1.1E-05	σ_{21}= 0.0E+00		
4	3.3072	σ_{11}= -1.41E-04	σ_{32}= -3.E-07	-0.14513	2.4
		σ_{22}= -1.50E-04	σ_{31}= 0.0E+00		
		σ_{33}= 4.1E-05	σ_{21}= 0.0E+00		
5	3.2873	σ_{11}= -1.77E-04	σ_{32}= 0.0E+00	-0.14060	2.9
		σ_{22}= -1.32E-04	σ_{31}= -1.9E-05		
		σ_{33}= -1.E-05	σ_{21}= 0.0E+00		

Table 4.12: In this table the final plane Cr distances are reported, the final stress tensors, the final pressures and the total energy differences with respect to the reference configuration for the different configurations.

Conf.	ΔE (eV)	D_x, D_y, D_z (10^{-4} Bohr^2)		
1	-3.46	1.07	0.59	13.64
2	-3.69	0.29	5.97	49.81
3	-3.86	0.16	1.28	39.52
4	-3.95	198.21	0.66	9.351
5	-3.82	0.71	111.40	5.661

Table 4.13: Bond energies and distortion/atom of C atoms (see Appendix A.2) in the three spatial directions.

4.7 Chromium intercalated between graphite planes

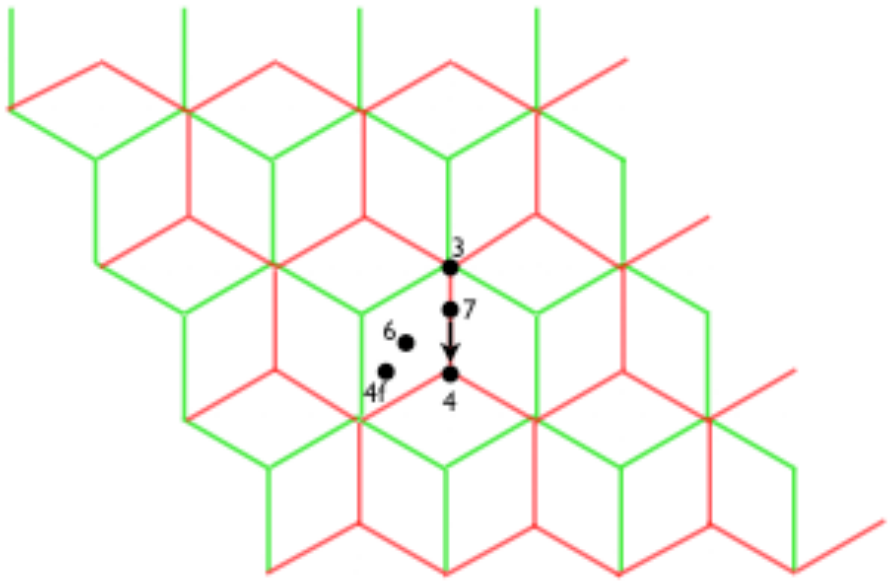


Figure 4.11: Trial configurations: in green the superior graphite plane, in red the inferior one. The arrow indicates the movement of the configuration 7 towards configuration 4 during geometry optimization. The 4f (four-fold) is the same configuration of the four-fold spiro interstitial defect studied with a C atom.

The trial configurations are shown in Fig. 4.11 and the cell is a 2x2x1 graphite cell with a Cr atom intercalated between planes. k-points grid is 2x2x2. The configuration 7 moves towards configuration 4 during geometry optimization and as can be seen in Tab. 4.14 the final energies are also similar. Configuration 3 produces a bound state with properties summarized in Tab. 4.14, 4.15.

Configurations 4f and 6 produce a shift of the graphite planes and the Cr ends up in the center of the hexagons up and down. This configuration is impossible in graphite where the planes are shifted one with the other but should be possible in carbon nanotubes. The shift of the planes is not physically relevant and is due to the wrong Van der Waals forces calculation in DFT. In fact all the XC energy functionals are of the short-range type Fig. 4.12. LDA

and GGA XC functionals have the following form:

$$E_{xc}^{LDA} = \int \epsilon_{xc}^{LDA}[n(\mathbf{r})]d^3\mathbf{r}$$
$$E_{xc}^{GGA} = \int \epsilon_{xc}^{GGA}\Big[n(\mathbf{r}), |\nabla n(\mathbf{r})|\Big]d^3\mathbf{r} \tag{4.21}$$

From these expressions is easy to understand that interaction is zero when the density is zero while Van der Waals forces are dispersion long-range forces. In some cases it is necessary to take correctly into account these interactions and one can add a classical potential which reproduces such forces [1, 75].

To calculate the bond energy of this configuration we have calculated the total energy of two overlapped graphene planes. This energy is 0.01 Ha greater than the total energy of the same number of atoms in graphite. This configuration has the lowest energy. The fact that this configuration is more bonded than others is consistent with the surface calculations. In fact in that case the minimum energy, among all trial configurations, was reached when the Cr atom is in the center of a graphene hexagon.

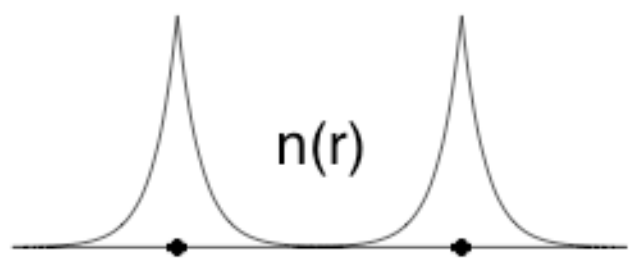


Figure 4.12: Short range interaction.

Conf.	Plane distance (Bohr)	Final Stress(Ha/Bohr^3)		ΔE (Ha)	Pressure (GPa)
3	3.15537	σ_{11}= -2.09E-04	σ_{32}= 1.7E-05	-0.17539	7.6
		σ_{22}= -2.57E-04	σ_{31}= 0.00E+00		
		σ_{33}= -3.13E-04	σ_{21}= 0.00E+00		
4	3.31363	σ_{11}= -2.74E-04	σ_{32}= -1.5E-05	-0.20400	7.8
		σ_{22}= -3.24E-04	σ_{31}= 0.00E+00		
		σ_{33}= -1.97E-04	σ_{21}= 0.00E+00		
4f	3.14578	σ_{11}= -2.60E-04	σ_{32}= 2.E-06	-0.25434	7.2
		σ_{22}= -3.06E-04	σ_{31}= -7.E-06		
		σ_{33}= -1.67E-04	σ_{21}= -3.9E-5		
6	3.15537	σ_{11}= -2.59E-04	σ_{32}= -5.E-07		7.1
		σ_{22}= -3.04E-04	σ_{31}= 7.E-07	-0.25438	
		σ_{33}= -1.63E-04	σ_{21}= -3.9E-05		
7	3.32234	σ_{11}= -2.68E-04	σ_{32}= -1.E-06	-0.20421	7.7
		σ_{22}= -3.23E-04	σ_{31}= 0.00E+00		
		σ_{33}= -1.91E-04	σ_{21}= 0.00E+00		

Table 4.14: In this table are reported final Cr distances from the lower plane, the final stress tensors, the final pressures and the total energy differences with respect to the reference configurations (energy of graphite plus isolated Cr atom for configurations 3-4-7 and energy of overlapped graphene planes plus isolated Cr atom for the configurations 4f-6) for the different configurations.

Conf.	ΔE (eV)	D_x, D_y, D_z (10^{-4} Bohr^2)		
3	-4.77	0.46	4.32	1393.63
4	-5.55	0.85	3.32	292.23
4f	-6.92	//	//	//
6	-6.92	//	//	//
7	-5.56	//	//	//

Table 4.15: Bond energies and distortion/atom of C atoms in the three spatial directions.

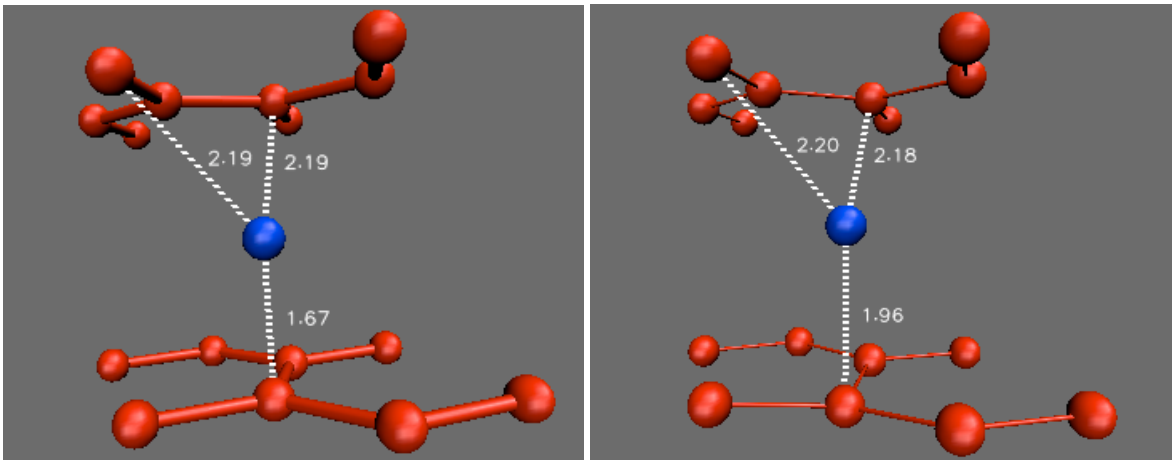
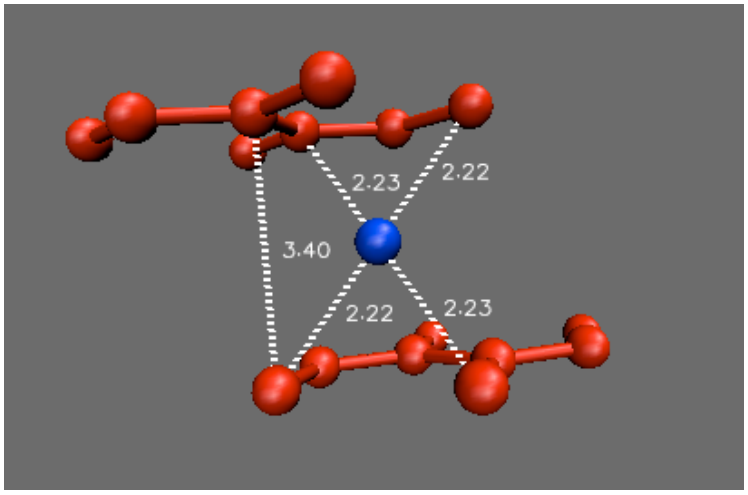


Figure 4.13: Above, the initial and final distances in Å between atoms in the configuration 4. Below, final distances in the configuration with shifted planes.



5 Calculations with ultrasoft pseudopotentials

5.1 Tests of pseudopotentials

From now on all calculations reported were done with CPMD code and ultrasoft library pseudopotentials (generated with "uspp" code) for Cr and C were used. As we saw, the cut-off needed for Chromium is higher with respect to Carbon. This is even more true in the following calculations, because Cr pseudoptential has also the $3s^2 3p^6$ as valence states. For this reason we only did a cut-off convergence study for Cr. All calculations were done in Γ point approximation since larger systems were considered. The exchange correlation functional is PBE of GGA type and the convergence criteria for forces are as those found in previous chapter.

The cut-off convergence study was done considering the Cr dimer Tab. 5.1. As can be seen in Tab. 5.1 bond length of Cr_2 converges much faster than energy. In general energy differences always need a higher cut-off than geometric properties or time averages of observables. A cut-off of 75 Ry was chosen for all total energy differences reported in this section. Even if 75 Ry is not so smaller than 100 Ry of the previous chapter, the ultrasoft Cr pseudopotential contains the $3s^2 3p^6$ as valence states. A similar choice for a norm-conserving pseudopotential increases the cut-off energy very much. The estimated cohe-

Cut-Off (Ry)	Δ E (eV)	Bond Length (Bohr)
50	0.00	3.2452
60	-0.67	3.2474
75	-0.25	3.2458
85	0.07	3.2460

Table 5.1: Cut-off convergence for a dimer of Chromium. ΔE is the atomization energy difference.

sive energy of bulk Chromium, calculated with a 4x4x4 cell and of Graphite with a 4x4x1 cell are:

$$E_{ch}^{C} = -7.60 \text{ eV}, \quad E_{ch}^{Cr} = -4.03 \text{ eV} \tag{5.1}$$

A Chromium dimer is a difficult molecule to treat in DFT and for an acceptable description semicore states must be included in the pseudopotential [6, 13]. Experimental results for bond length and atomization energy of Cr_2 are respectively:

$$d_{Cr_2} \simeq 3.2 \text{ Bohr}, \quad E_{atomiz} \simeq 1.5 \text{ eV} \tag{5.2}$$

Our results for a cell of $\simeq$10 Å are reported in Tab. 5.2 The bond length has been well estimated without spin but not with it. Conversely atomization energy is better estimated with spin than without it. This situation is characteristic of other transition metals like Cr (see [6] and references therein). In [13] atomization energy and bond length were estimated using the Multireference Configuration Interaction method and was found that:

$$d_{Cr_2} = 3.25 \text{ Bohr}, \quad E_{atomiz} = 1.09 \text{ eV} \tag{5.3}$$

even with this too expansive method the atomization energy is quite far from its experimental value and similar to ours. Conversely the bond length is better estimated with the Multireference Configuration Interaction method . This method however can be adopted only to treat few atoms. Total energies

Spin	Atomization energy (eV)	Bond Length (Bohr)
No	10.92	3.22
Yes	1.07	4.33

Table 5.2: Final results of the Cr dimer.

differences of configuration 4 and with shifted planes, reported in section 4.7, were recalculated with ultrasoft pseudopotentials in a 4x4x1 cell. The results are reported in Tab. 5.3.

	CPMD	Abinit
Conf	Δ E(eV)	Δ E(eV)
4	-5.09	-5.55
Shift	-5.97	-6.92

Table 5.3: Bond energies of configurations 4 and with shifted planes calculated with Abinit and CPMD. Note that also cell dimensions are different: 2x2x1 with Abinit and 4x4x1 with CPMD.

5.2 Graphene surface with vacancy

As we saw the Cr bind to a perfect graphene surface. The same type of calculations were done with a graphene surface with a vacancy. At first the formation energy of a vacancy was calculated with a 4x4x1 cell:

$$E_{form} = 7.66 \text{ eV} \tag{5.4}$$

This value is in perfect accordance with previous calculations (see [40] and references therein). In all calculations reported in this section the cell is 2x2x1 (with only the plane at $z = 0$) and one atom of graphene plane is fixed to avoid plain from shifting rigidly during geometry optimization. The trial configurations are represented in Fig. 5.1. The configuration Pushed 1 is the same as configuration 1 but in this case the Cr atom is put at a zero initial distance from the plane and its z-coordinate is fixed during optimization.

The atom of Cr in the configuration 3 moves towards the position of Cr in the configuration 1 during geometry optimization. Configurations 1 and 2 produce two bound states Tab. 5.5. The distances between atoms after geometry optimization are reported in Fig. 5.2-5.3.

The configuration pushed 1 was considered to try to understand if Cr atom can penetrate the graphene plane through vacant site. The results of calculations suggest that the Cr atom should not penetrate the plane so easily. In fact in configuration pushed 1 the graphene plane deforms to reproduce the same type of bounds of configuration 1. As can be seen in Tab. 5.4 energy contributions of the two configurations are quite similar except that in configuration 1 the ΔLPs (long-range) contribution is larger that $\Delta NLPs$ (short-range) while in configuration Pushed 1 the contrary happens. From Tab. 5.4 it is clear that the plane distorts itself more than in configuration 1 along the z-axis while distortions along other directions are similar. See also Fig. 5.2.

Conf.	ΔEs (Ha)	ΔLPs (Ha)	$\Delta NLPs$ (Ha)	ΔXC (Ha)	ΔE_{GGA} (Ha)
1	-0.10049	-0.94012	-0.19744	-0.53677	0.05161
2	-0.17111	-0.49466	-0.10743	-0.41496	0.04806
Pushed 1	-0.12227	-0.77642	-0.53641	-0.50862	0.05802

Table 5.4: Differences of some energy contributions with respect to reference configuration.

Conf.	ΔE (eV)	D_x, D_y, D_z		10^{-3} Bohr^2	Distance (Bohr)
1	-9.77	2.544	1.264	74.128	2.5396
2	-6.40	2.691	3.629	64.170	2.9326
Pushed 1	-8.81	2.637	9.243	967.390	0.0000

Table 5.5: The bond energy ΔE is the total energy difference with respect to the energy of a graphene surface with a vacancy plus the total energy of an isolated Cr atom. The distance is the final z-coordinate of Cr atom.

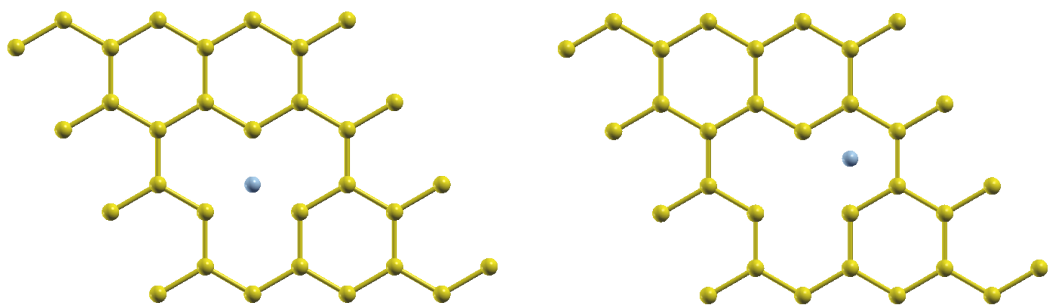
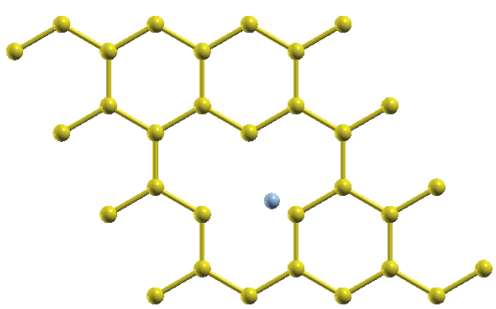


Figure 5.1: Above on the left the relative position of the Cr atom with respect to the graphene surface in the configuration 1; on the right in the configuration 2. Below in the configuration 3.



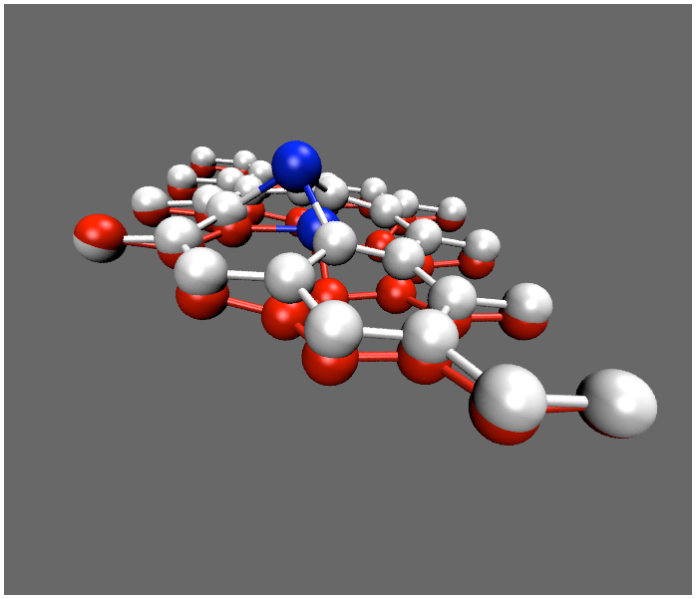
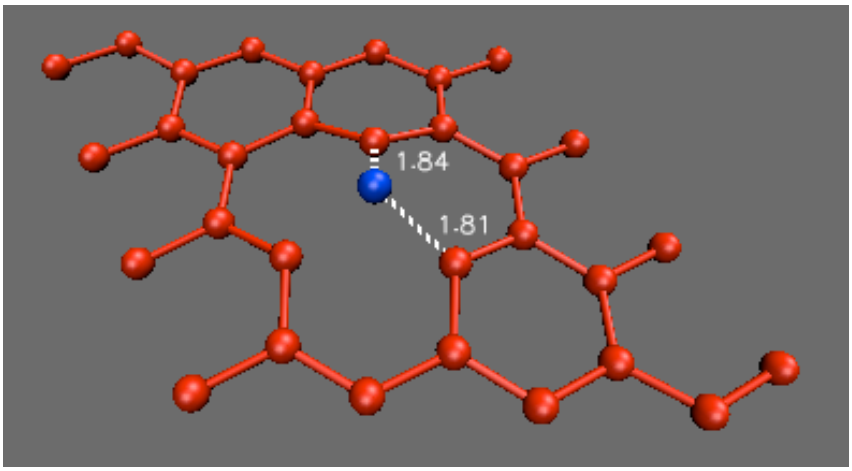


Figure 5.2: Above, the two configurations one. Below, inter atomic distances in Å after geometry optimization in the configuration 1.



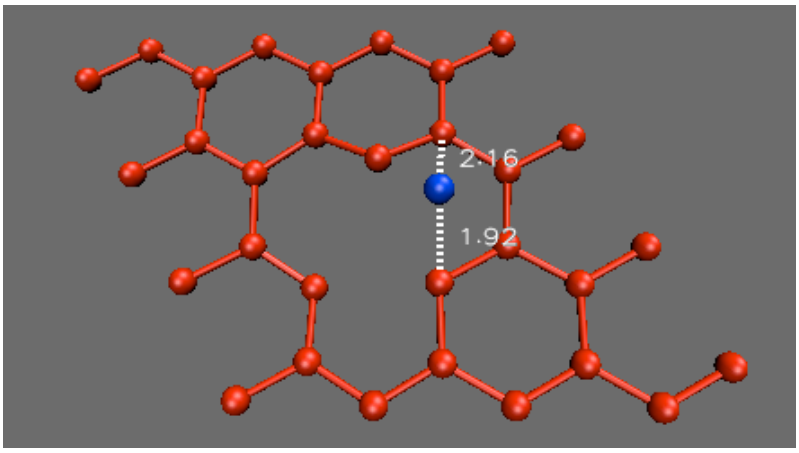
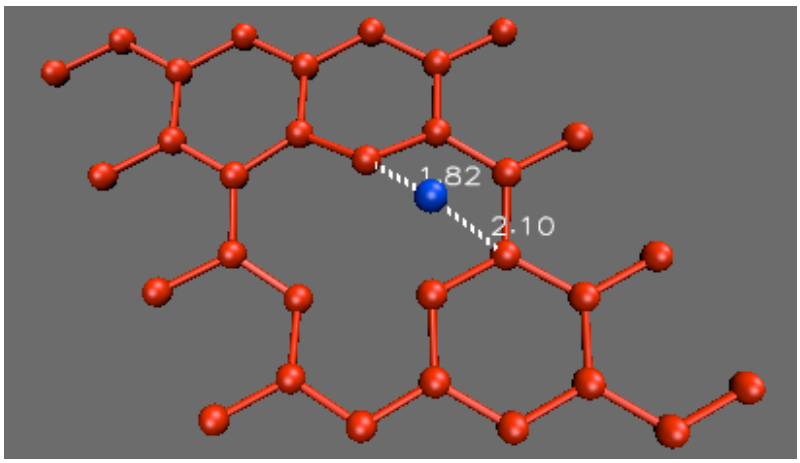


Figure 5.3: Interatomic distances in Å after geometry optimization in the configuration 2.



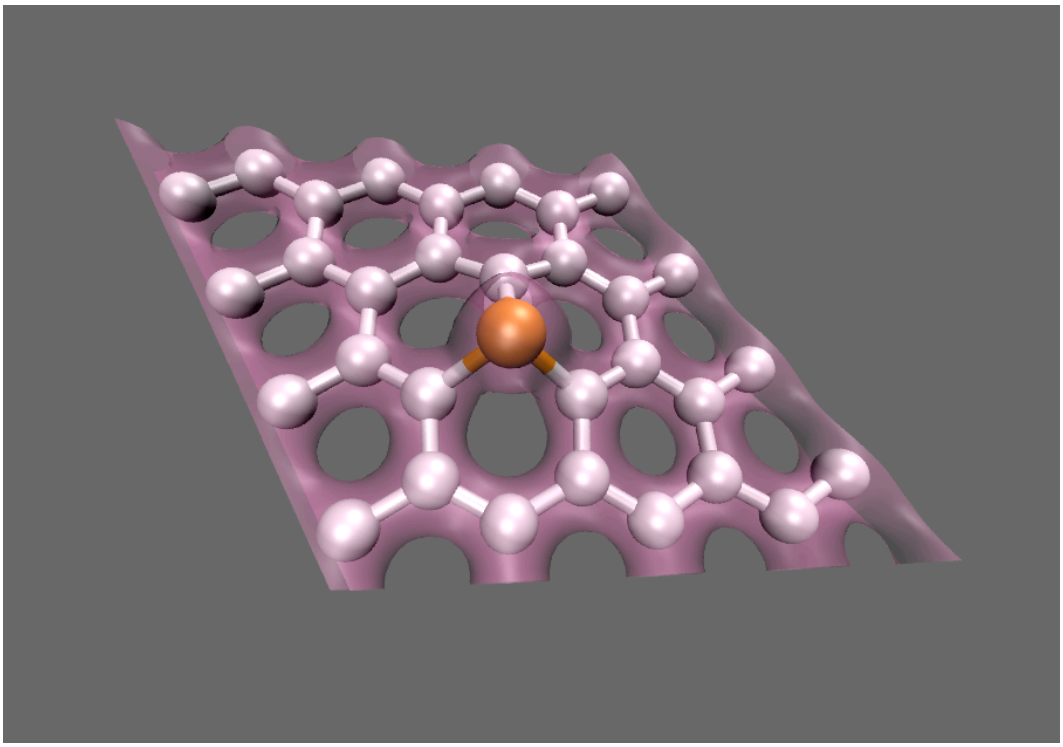
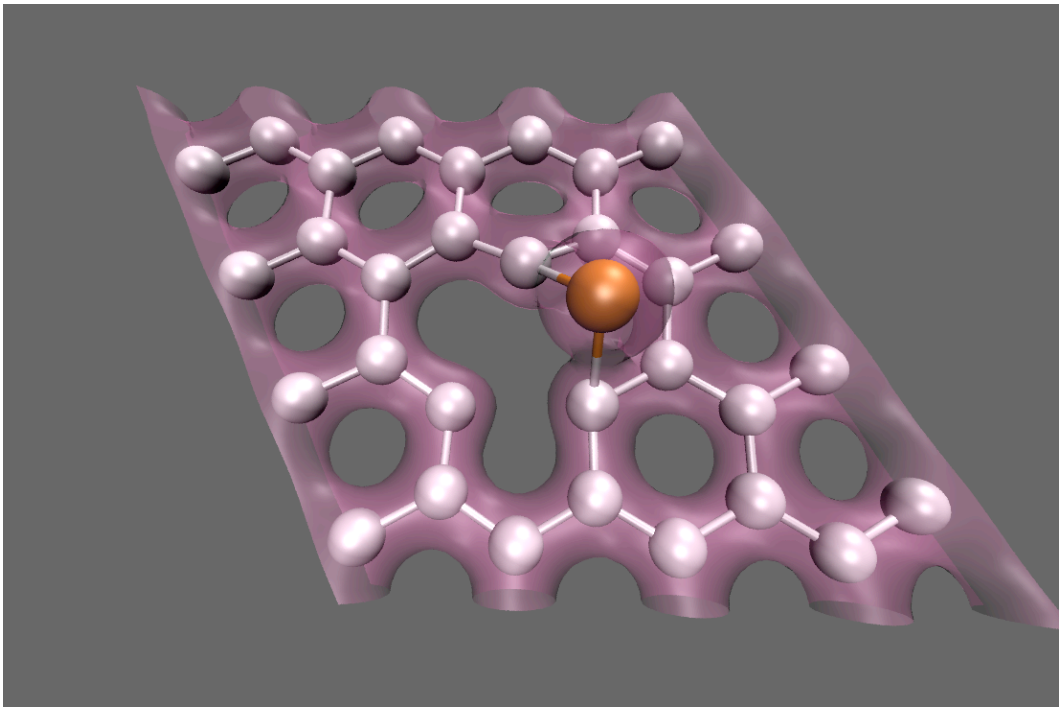


Figure 5.4: Isodensity surface for $\rho \simeq 0.1$ in presence of vacancy. Above, for configuration 1, below, for configuration 2.



5.3 Chromium intercalated in graphite with vacancy

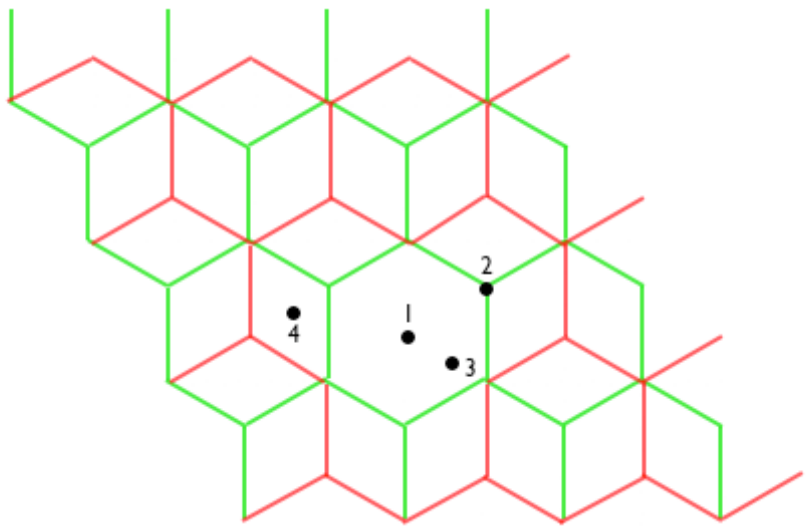


Figure 5.5: Initial positions of the atoms in the four trial configurations.

In this section the results of adhesive properties of bulk graphite with a vacant site and Cr atom are reported. The trial initial positions of atoms are reported in Fig. 5.5 and the cell is 4x4x1 as before. Configurations 1 and 2 produce two bound states (Tab. 5.6, Fig. 5.6). The Cr atom in configuration 3 moves toward the position of the Cr atom in the configuration 1. Configuration 4 produces an unphysical shift of the graphite planes as happened for graphite planes without vacancy with norm-conserving pseudopotentials.

In this and the previous section configurations with very large bound energies were found ($\simeq$ 10 eV). On this purpose, we can note that separations between states $(3s, 3p)$, not pseudized with norm-conserving pseudopotentials, and states $(3d, 4s)$ is not so away from bound energies found Tab. 5.8. Then a characterization of these configurations with norm-conserving pseudopotentials should result wrong.

Conf.	ΔE (eV)	D_x, D_y, D_z 10^{-3} Bohr^2			Distance (Bohr)
1	-10.79	0.962	2.431	29.984	2.6875
2	-7.77	31.097	4.224	17.051	2.8229

Table 5.6: The ΔE is the total energy difference with respect to the energy of bulk graphite with a vacancy plus the total energy of an isolated Cr atom. The distance is with respect to the lower plane.

Conf.	ΔEs (Ha)	ΔLPs (Ha)	$\Delta NLPs$ (Ha)	ΔXC (Ha)	ΔE_{GGA} (Ha)
1	-0.12144	-1.15278	-0.08024	-0.67495	0.07067
2	-0.13671	-0.87269	-0.02895	-0.58376	0.07500

Table 5.7: Differences of energy contributions with respect to reference configuration.

State	Occupation	ΔE (eV)
$3s$	2	0.00
$3p$	6	28.71
$3d$	1	44.35
$3s$	5	1.09

Table 5.8: Energy separation of Cr levels for pseudized orbitals.

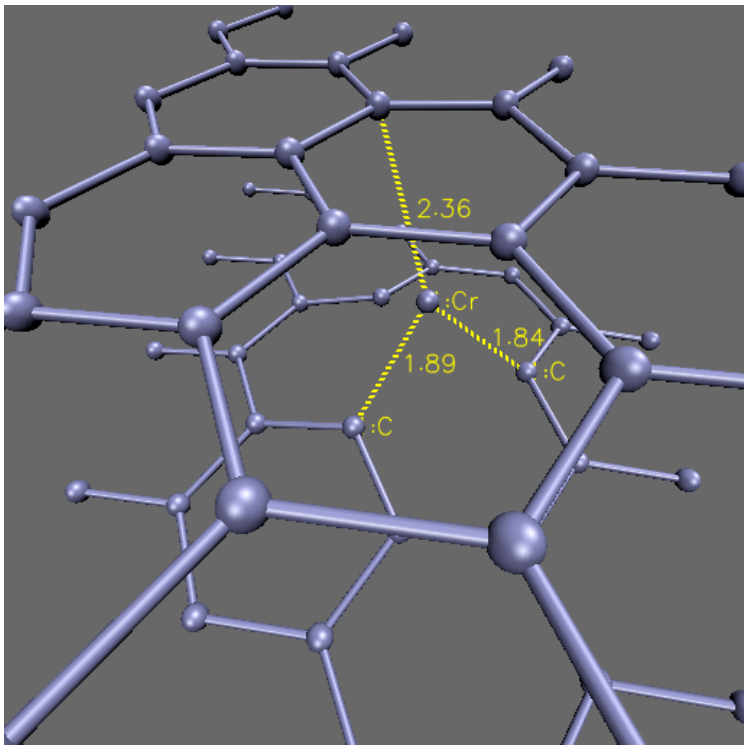
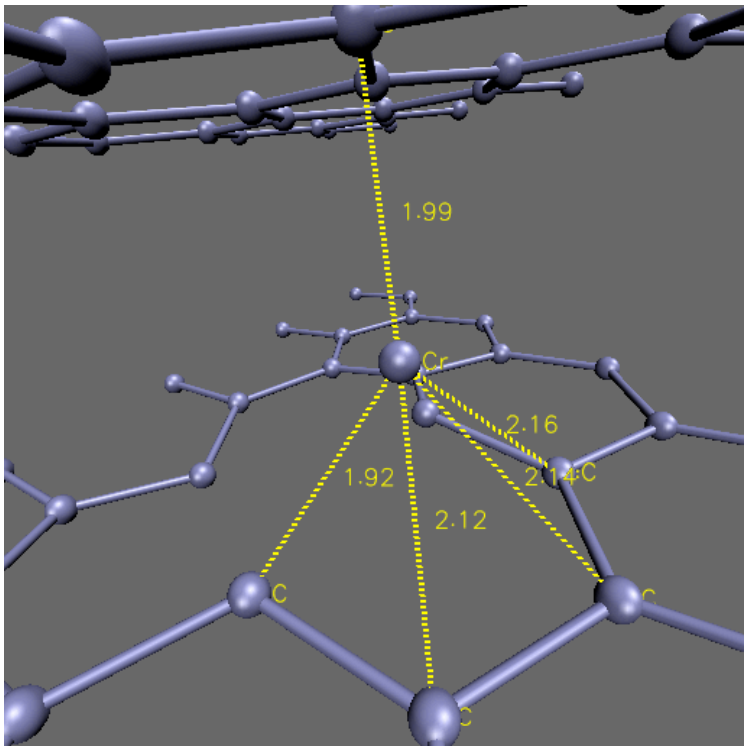


Figure 5.6: Above, final inter-atomic distances in Å in configuration 1, below, in configuration 2.



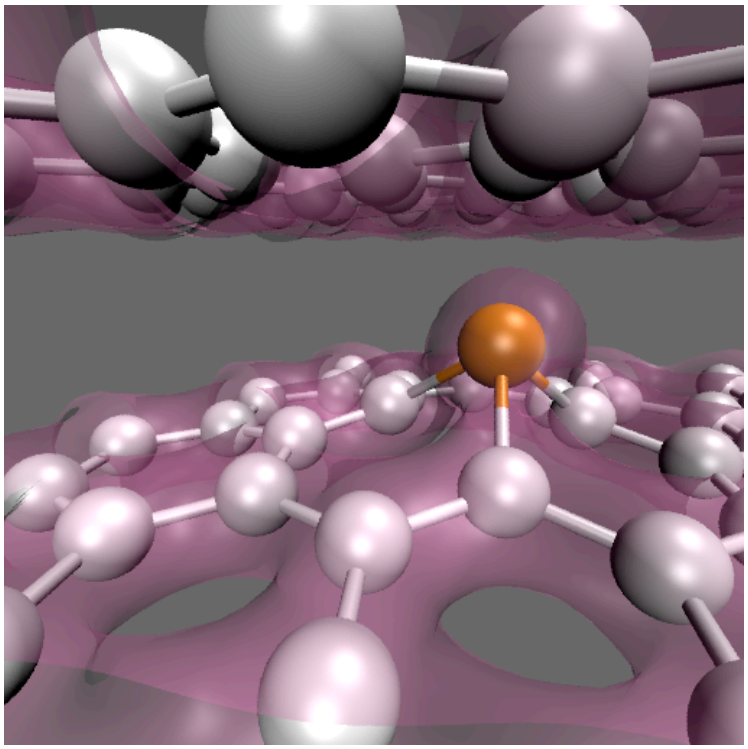
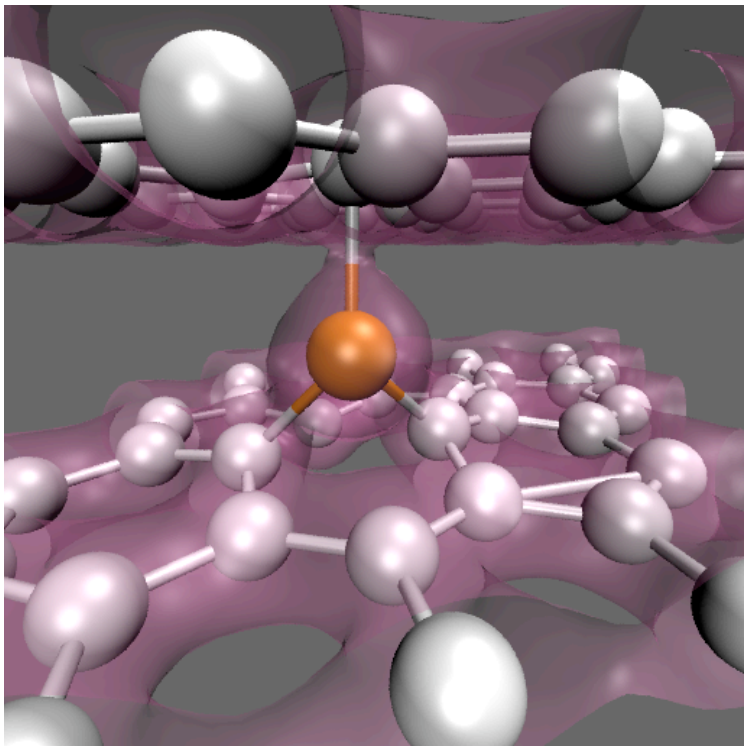


Figure 5.7: Isodensity surfaces for $\rho \simeq 0.1$. Above, for configuration 1, below, for configuration 2.



5.4 Graphene surfaces and Cr_2

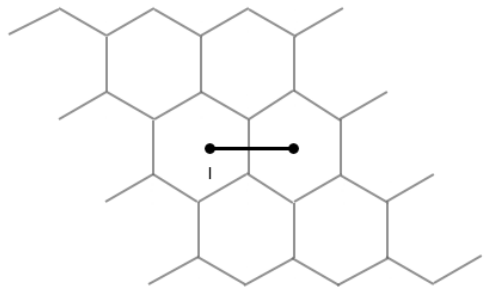


Figure 5.8: Positions relative to the graphene surface in configuration 1

We have also studied the bonding properties of a Cr_2 molecule on a perfect graphene surface and on a graphene plane with a vacant site or a SW defect. In the case of perfect plane the cell used is 2x2x2 (with only the plane at $z = 0$), that is $\simeq$ 24 Bohr in z direction. In fact the bond lengths in this case are larger and a bigger cell in z direction has to reduce the effect of periodic boundary conditions. Only the lower graphene plane at zero z coordinate is present. Trial configuration 1 is showed in Fig. 5.8. In the configuration 2 the Cr_2 dimer is perpendicular to the graphene surface and the lowest Cr is above the center of an hexagon of C atoms (Fig. 5.9). In the Tab. 5.9, 5.10 calculations details are reported.

Conf.	ΔE (eV)	D_x, D_y, D_z	10^{-3} Bohr^2		Distance (Bohr)
1	-0.07	0.762	16.645	0.050	6.5339
2	-0.30	0.070	0.606	4.334	3.7035

Table 5.9: Results of calculations in the configurations 1 and 2. Distances refer to the nearest plane Cr atom.

Conf.	ΔEs (Ha)	ΔLPs (Ha)	$\Delta NLPs$ (Ha)	ΔXC (Ha)	ΔE_{GGA} (Ha)
1	0.00249	0.00613	0.00366	-0.00186	0.00559
2	-0.13594	0.23699	-0.13593	-0.02860	0.02485

Table 5.10: Differences in energy contributions with respect to reference configuration.

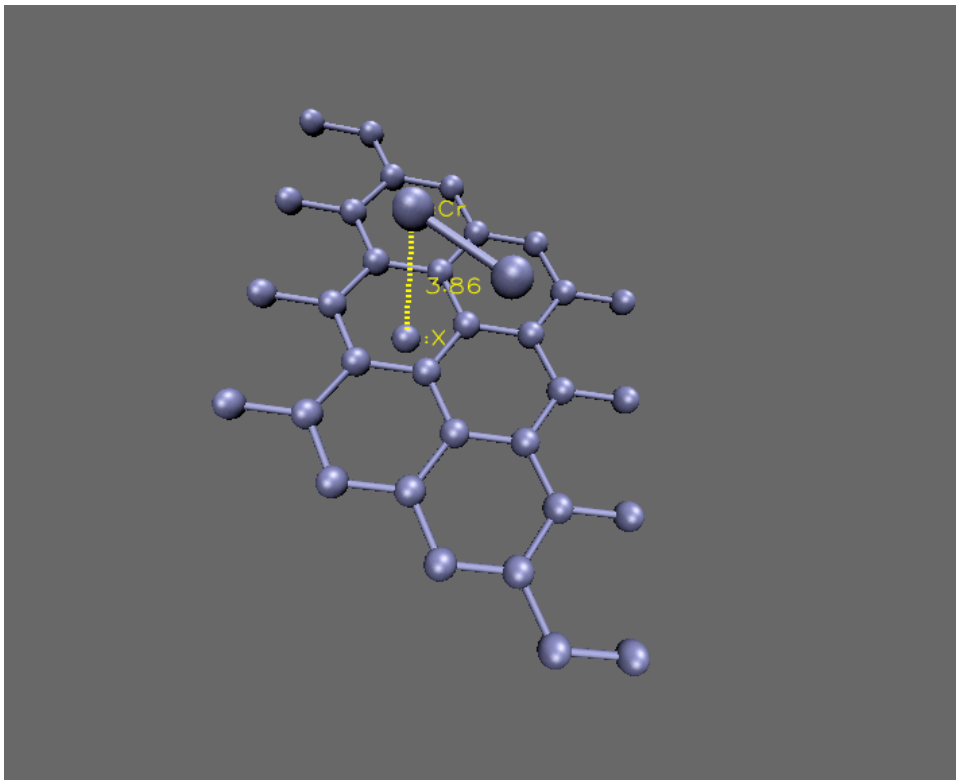
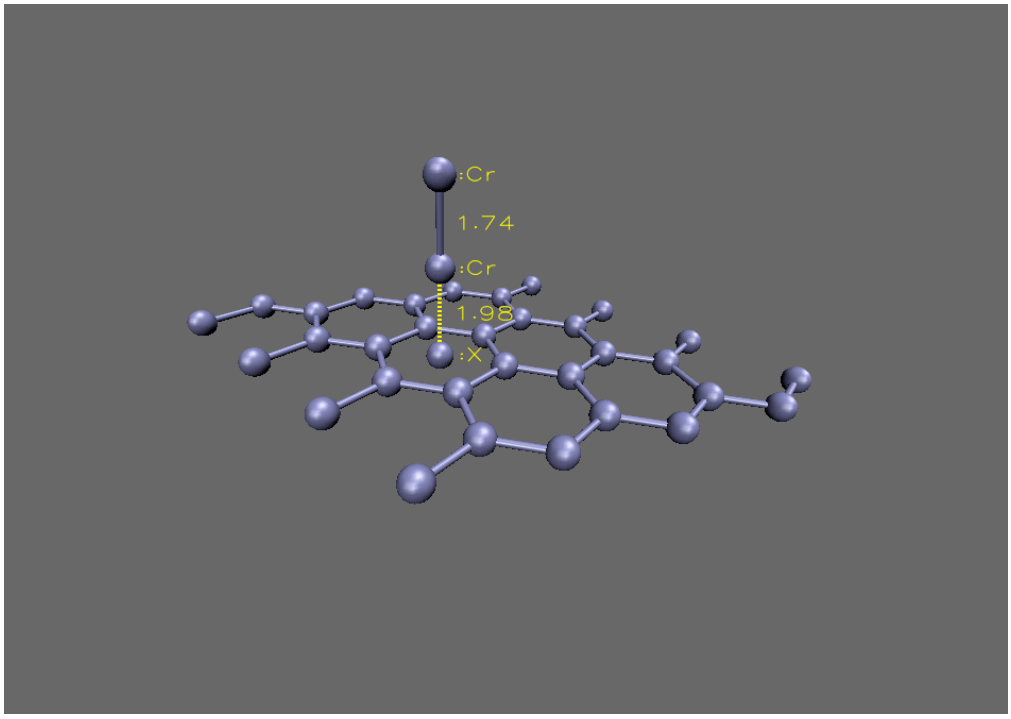


Figure 5.9: Above, final plane dimer distances in Å in configuration 1, below, in configuration 2.



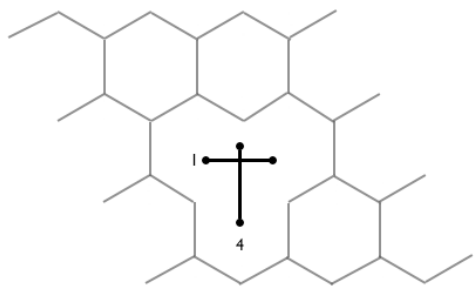


Figure 5.10: Positions of the Cr atoms in the configurations 1 and 4 relative to the graphene surface.

The bond energy of a Cr_2 dimer with a graphene surface with a vacancy are much higher than those with a perfect graphene plane. The cell used is 4x4x1.5, that is $\simeq$ 18 Bohr in z direction. Four trial configurations were chosen. Configurations 1 and 4 are shown in Fig. 5.10. In the configurations 2 and 3 the Cr_2 dimer is initially perpendicular to the plane and the lowest Cr atom is respectively in the position 1 and 2 of the Fig. 5.1. In these two last configurations the dimer begins to rotate (Fig. 5.12). This intermediate configuration is reconsidered next. The details of calculations of configurations 1 and 4 are shown in Tab. 5.11, 5.12.

Conf.	ΔE (eV)	D_x, D_y, D_z	10^{-3} Bohr^2		Distance (Bohr)
1	-3.72	1.548	14.153	70.158	3.6821
4	-4.68	2.914	3.156	113.462	2.7431

Table 5.11: Results of the calculations in the configurations 1 and 4.

Conf.	ΔEs (Ha)	ΔLPs (Ha)	$\Delta NLPs$ (Ha)	ΔXC (Ha)	ΔE_{GGA} (Ha)
1	-0.34622	-0.19734	-0.68328	-0.40082	0.03976
4	-0.43949	-0.20388	-0.73265	-0.49936	0.04095

Table 5.12: Differences in energy contributions relative to reference configuration.

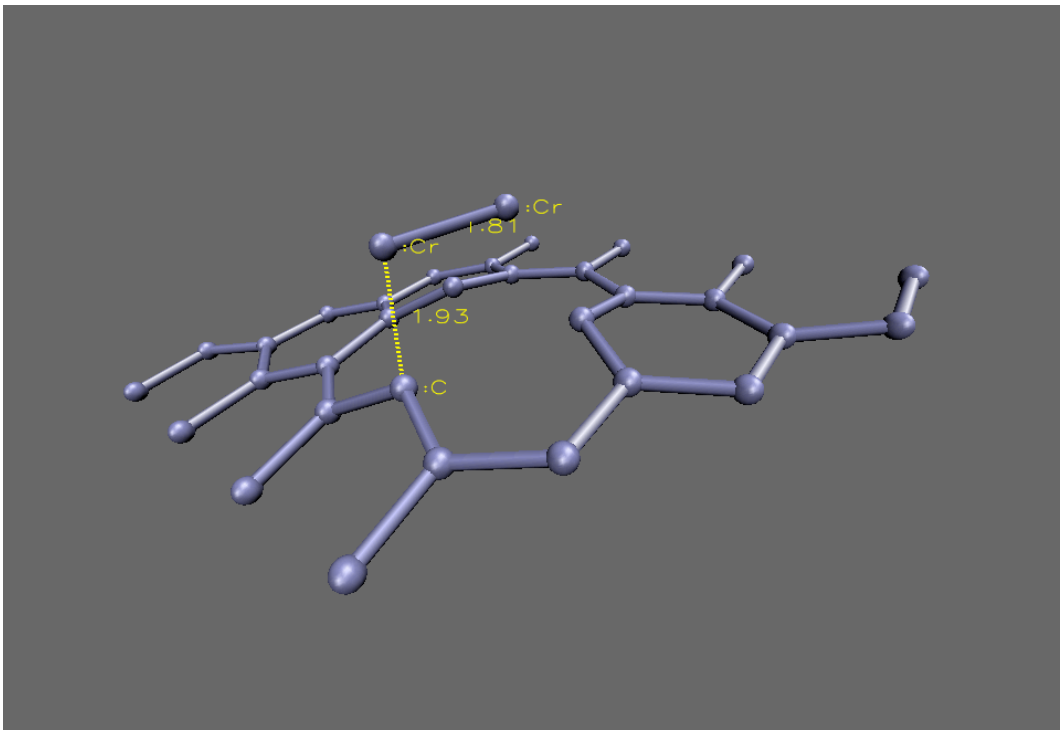
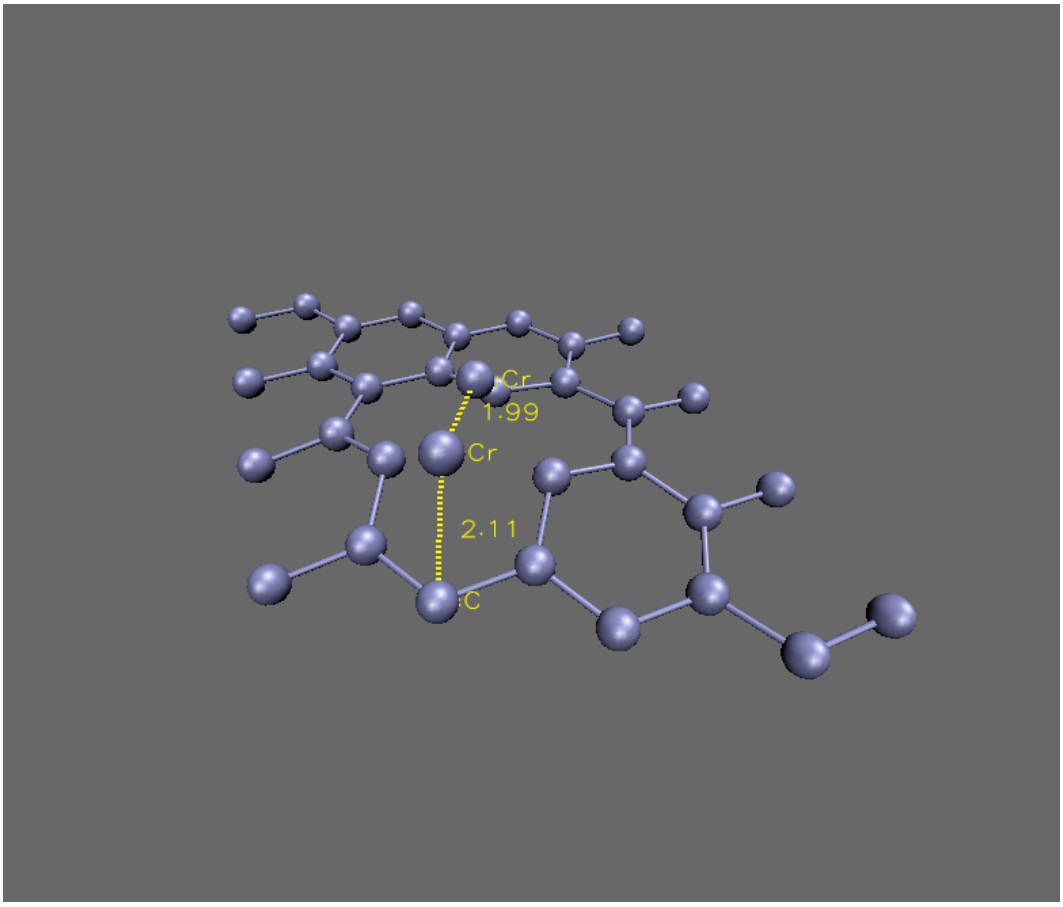


Figure 5.11: Final inter-atomic distances in Å for configuration 1.



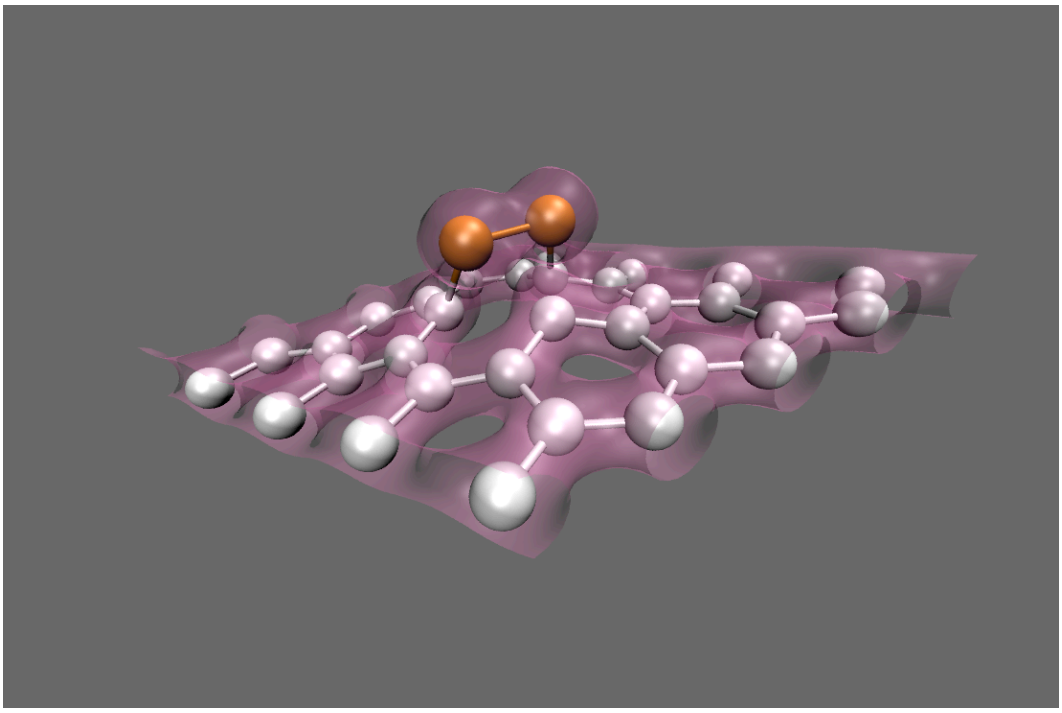
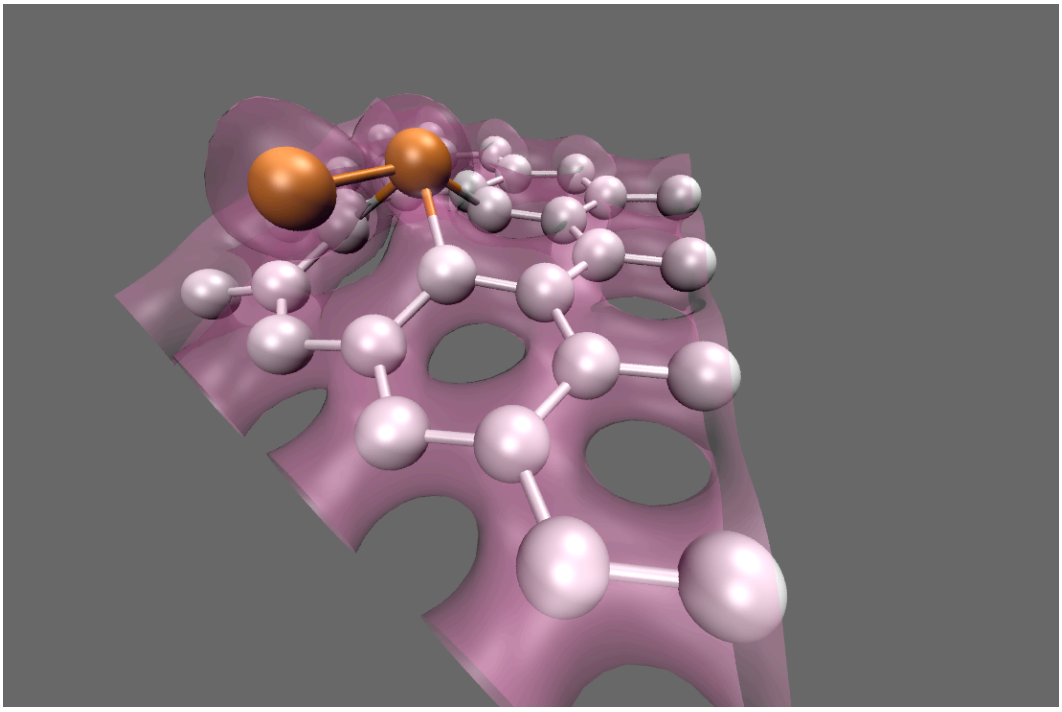


Figure 5.12: Above, isodensity surface for $\rho \simeq 0.1$ in configuration 1. Below, isodensity surface for $\rho \simeq 0.1$ in configuration 4.



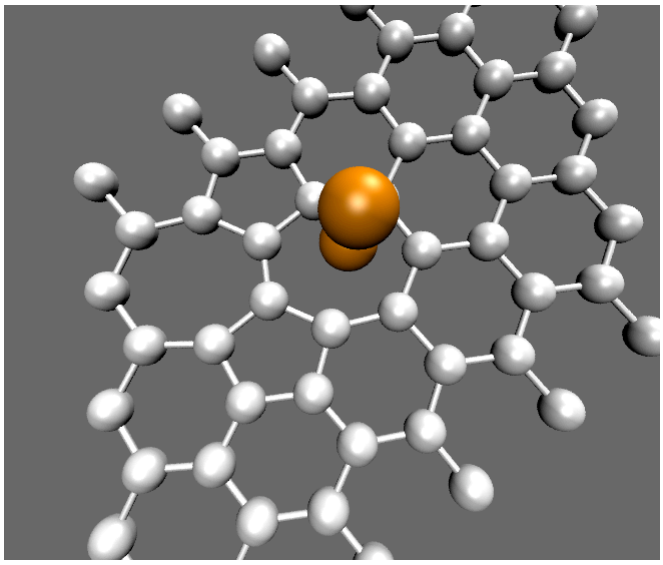


Figure 5.13: Position with respect to the graphene surface in the case of SW defect.

As we said previously if initially the dimer is perpendicular to the graphene plane it begins to rotate to align itself to the plane. The minimum configuration reached starting from these initial conditions is lower in energy with respect to the previously considered configurations. The same type of calculation was also done for a plane with an SW defect with the dimer centered on the 7 ring Fig 5.13. The energy formation of a 5x7 SW defect was estimated using a 4x5x1 cell and the result is:

$$E_{form} = 6.73 \text{ eV} \tag{5.5}$$

This value is in accordance with other DFT calculations (see [38] and references therein). In Tab. 5.13, 5.14 the results obtained starting with dimer perpendicular to the plane at a distance of 3 Bohr are summarized.

Type defect	ΔE (eV)	D_x, D_y, D_z		10^{-3} Bohr^2	Distance (Bohr)
Vacancy	-4.90	1.195	3.623	105.526	2.8026
SW	-0.74	0.057	0.087	1.695	3.2412

Table 5.13: Results of the calculations for two types of defect starting with a perpendicular dimer. Distances refer to the nearest plane Cr atom.

Defect type	ΔEs (Ha)	ΔLPs (Ha)	$\Delta NLPs$ (Ha)	ΔXC (Ha)	ΔE_{GGA} (Ha)
Vacancy	-0.45906	-0.16975	-0.19546	-0.43140	0.03967
SW	-0.19300	0.31757	-0.18155	-0.03343	0.02764

Table 5.14: Differences in energy contributions relative to reference configuration for two type of defect starting with a perpendicular dimer.

From these tables it is clear that in the case of a vacancy defect the dimer is more bounded to the plane. With respect to the minimum configuration 4 an extra charge localization between a Cr and a C atom can be noted Fig. 5.14. This should be the reason for the larger bond energy of this configuration with respect to configuration 4. For the case of SW defect the plane is distorted very little and even the bond energy is smaller. Moreover the dimer remains almost perpendicular to the plane Fig. 5.13.

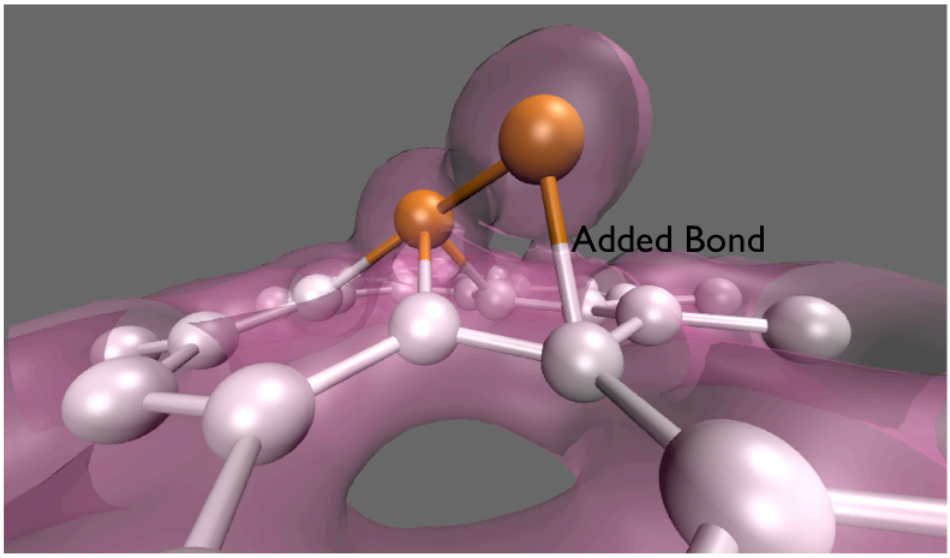
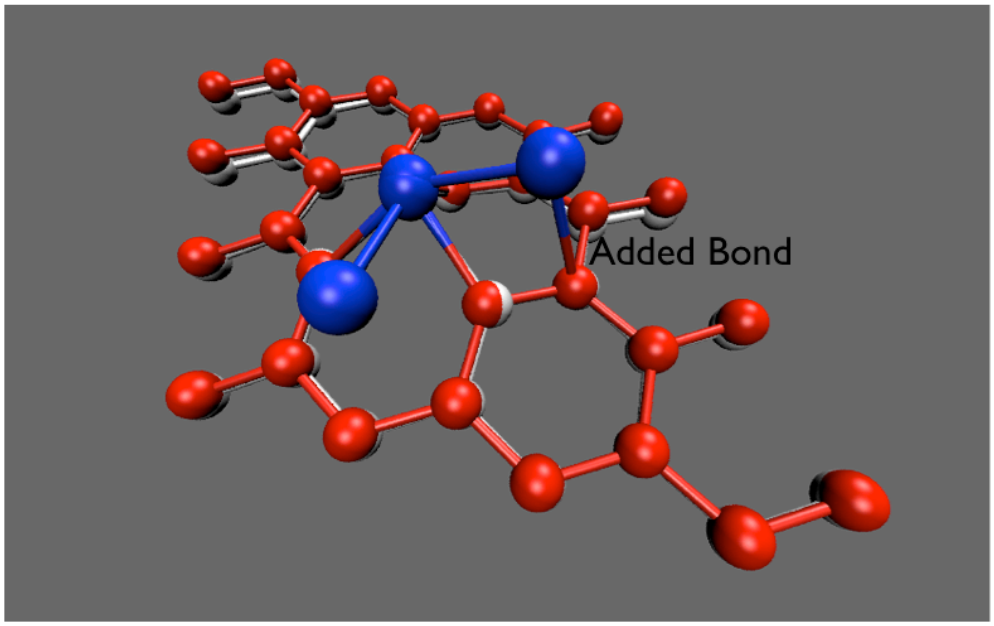


Figure 5.14: Above, the final configuration reached starting with dimer perpendicular to the surface, the extra bond present in this case is evident. Below, differences between the final position of the two configurations.



5.5 Formation and dissociation of a Cr dimer: total energies

In section 5.2 we have seen that for a Cr atom it should be very difficult to penetrate a graphene plane through a vacant site and it should be even more difficult to penetrate a perfect plane so we have tried to find a process in which the penetration process include a Cr dimer. Considering this process two problems were aroused. The first problem regards the possible presence of Cr dimers in the gas evaporated on outer shell of MWCNTs. In experimental situations we are interested in, a Cr gas is evaporated at a temperature of about 1500 K. This temperature corresponds to an energy of about 0.3 eV. Such temperature should not be sufficient to form dimers. However some dimers should be present in the evaporated gas. Then we investigated if two Cr atoms on a graphene plane initially placed at large distance (with respect to dimer bound distance $\simeq 1.7$ Å) tend to move further away or to approach each other. These however are preliminary calculations where convergence was not reached. A similar process was studied in higher detail in the free energy calculations reported in the next section.

The trial configurations are shown in Fig. 5.15. Starting from the configuration 1 and 2 the distance between Cr atoms becomes much smaller than the initial ones (Fig. 5.16).

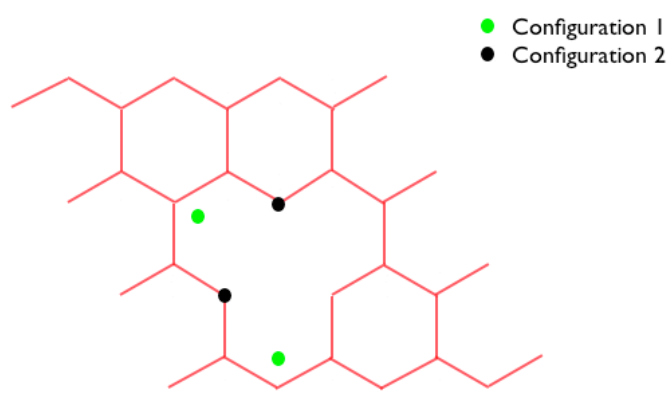


Figure 5.15: Trial configurations of two Cr atoms on a graphene surface with a vacancy.

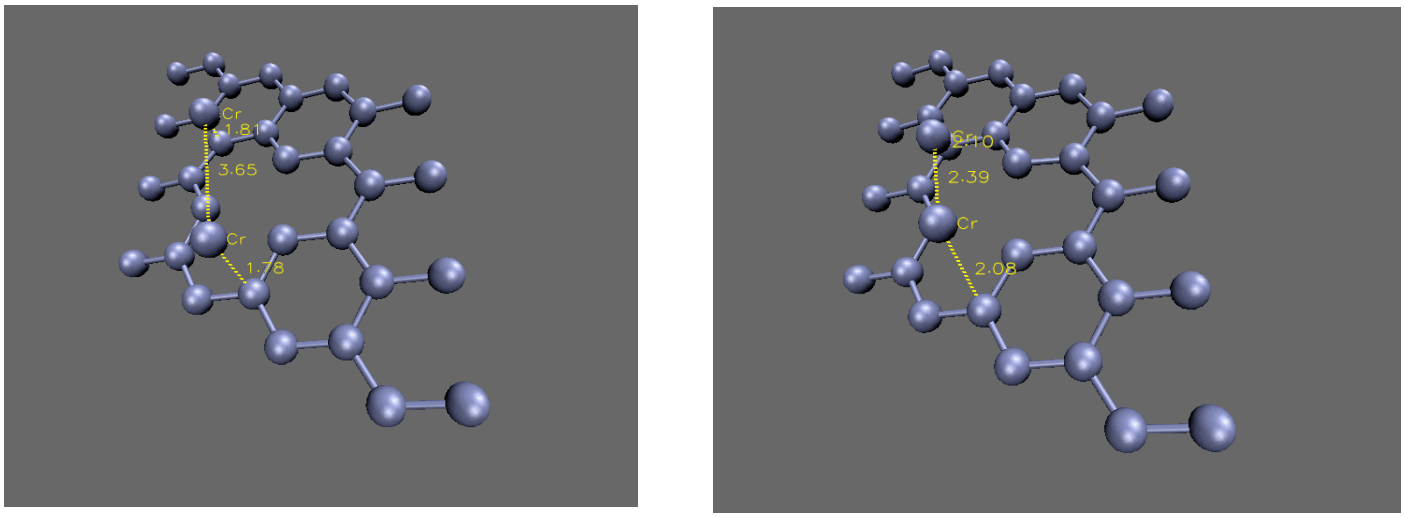
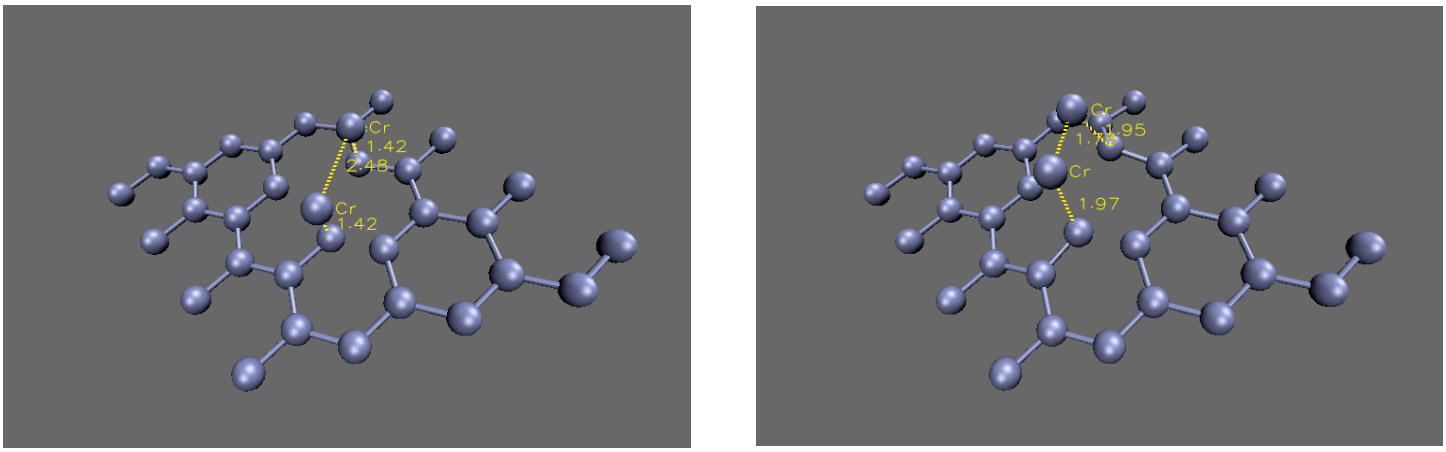


Figure 5.16: Above initial and final distances in Å for configuration 1. Below for configuration 2.



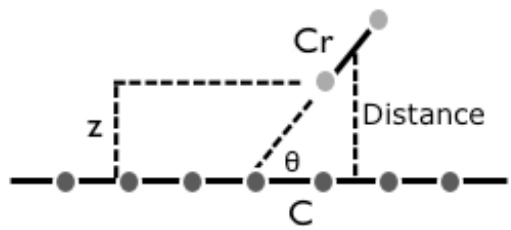


Figure 5.17: Illustrations of z, θ and plane-dimer distance.

These calculations suggest that the dimer formation on a defective graphene plane could be possible.

The second problem regards the dissociation process itself. In a trial MD run for a dimer initially not stretched through a vacant site, it was found that dimer practically dissociates in few MD steps at 350 K. These two facts suggest to search for a process in which the dimer penetrates the plane and dissociates.

For these reasons was studied the energy landscape of a Cr dimer which is placed closer and closer to the ideal plane situated at zero z coordinate. Calculations were done both for a SW-defective plane and for a vacancy-defective plane. In these calculations a C atom of the plane is kept fixed to avoid a shift of the plane itself. The dimer is instead placed on the plane keeping fixed only the z coordinate of the initial nearest plane Cr atom Fig. 5.17. In the first and the last configurations in Tab. 5.15 and Fig. 5.18 , the dimer is completely free so these are local minima Fig. 5.14,5.13. In general even if the dimer is initially perpendicular to the plane during geometry optimization, it rotates and the angle to the ideal plane is called θ Fig. 5.17. The plane dimer distance is defined as the distance between the ideal plane and the dimer center Fig. 5.17. Results are summarized in Tab. 5.15, Fig. 5.18.

In the case of a vacancy defect a 4x4x1.5 cell (with only a plane at $z = 0$) was used. The more stretched configuration is also the more stable one and the total energy barrier to reach such configuration is 1.3 eV.

In the case of a Stone-Wales defect a 5x4x1.5 (with only a plane at $z = 0$) cell was used. The more stretched configuration is not the more stable. If a calculation is started from the more penetrating configuration, adding the inferior perfect graphene plane and the dimer completely free, a state with binding energy of -1.12 eV and a stretch of 0.55 Å is reached. Also this

configuration is not the most stable one. Moreover from Fig. 5.20 it is clear that the penetrating dimer tends to pull out a carbon atom so as to prefer a vacancy.

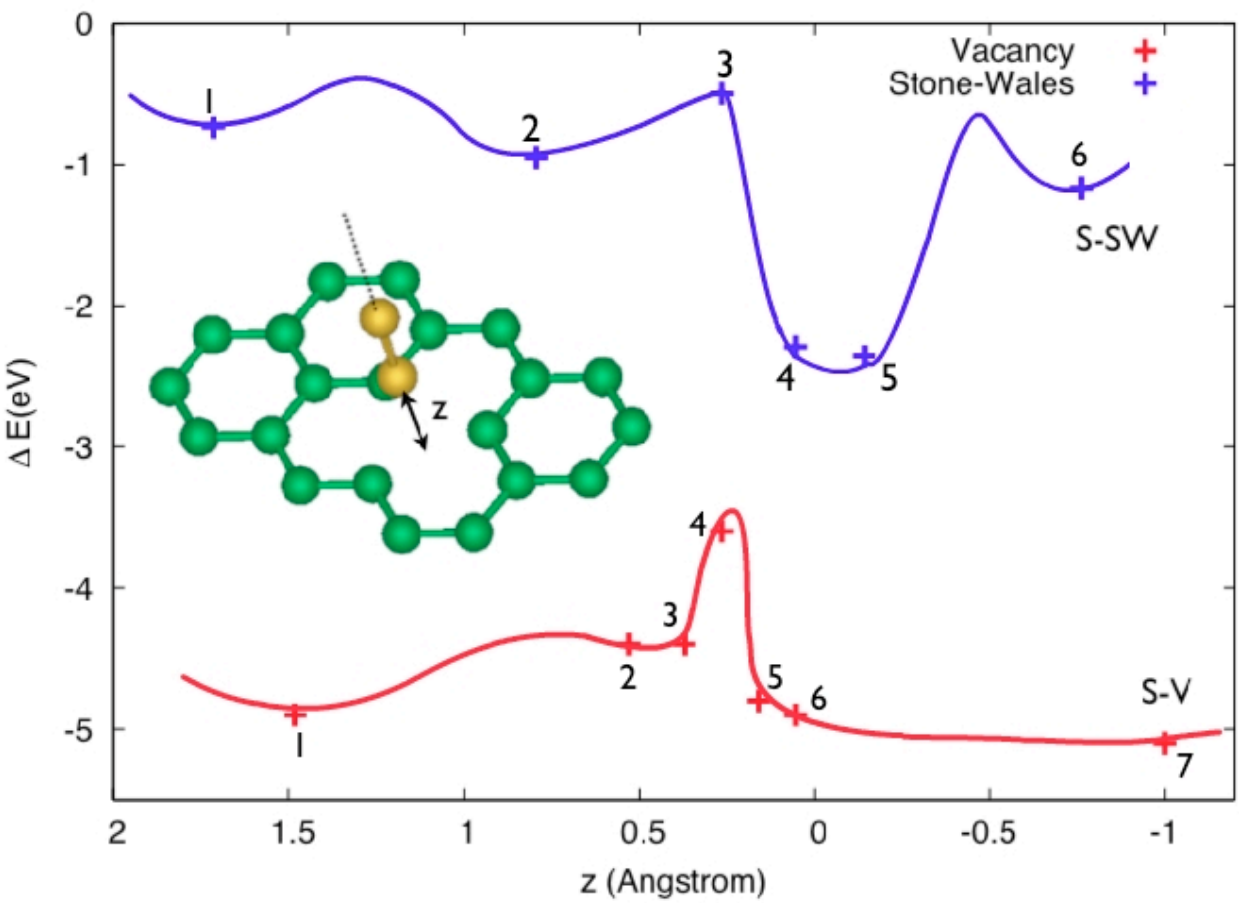


Figure 5.18: Energy landscape in function of the dimer distance z (Fig. 5.17). These curves are symmetric around the points labeled in figure as S-SW and S-V.

Vacancy

z (Å)	Distance (Å)	θ^o	Stretch (Å)	ΔE (eV)
1.48	1.81	19.2	0.28	-4.9
0.53	1.21	43.4	0.28	-4.4
0.37	1.20	59.8	0.22	-4.4
0.26	0.82	34.1	0.30	-3.6
0.16	1.00	59.8	0.30	-4.8
0.05	0.90	60.2	0.24	-4.9
-1.00	0.00	58.9	0.30	-5.1

Table 5.15: Energy landscape for the two type of defects. z is the z coordinate of initially nearest Cr atom. The stretch refers to Cr dimer. ΔE is the binding energy.

Stone-Wales

z (Å)	Distance (Å)	θ^o	Stretch (Å)	ΔE (eV)
1.71	2.60	90.0	0.05	-0.74
0.80	1.64	72.9	0.07	-0.95
0.26	1.12	71.9	0.09	-0.50
0.05	0.86	54.5	0.28	-2.29
-0.14	0.66	53.5	0.30	-2.36
-0.76	0.20	67.9	0.39	-1.17

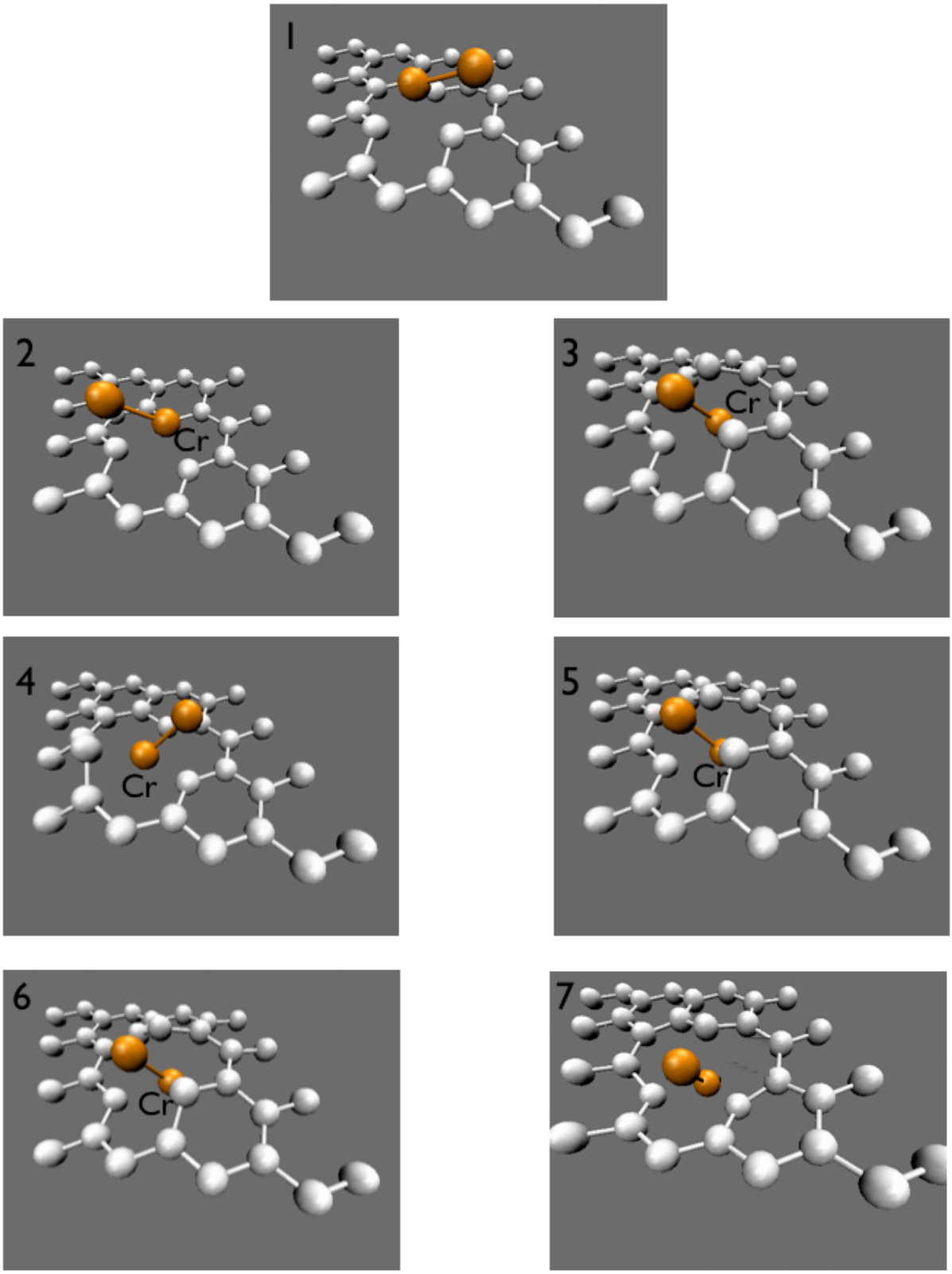


Figure 5.19: Optimized geometries in the configurations 1,2,3,4,5,6,7 in the case of vacancy defect (see Fig. 5.18). The z coordinate of the Cr atoms labeled in figures was kept fixed during the optimizations.

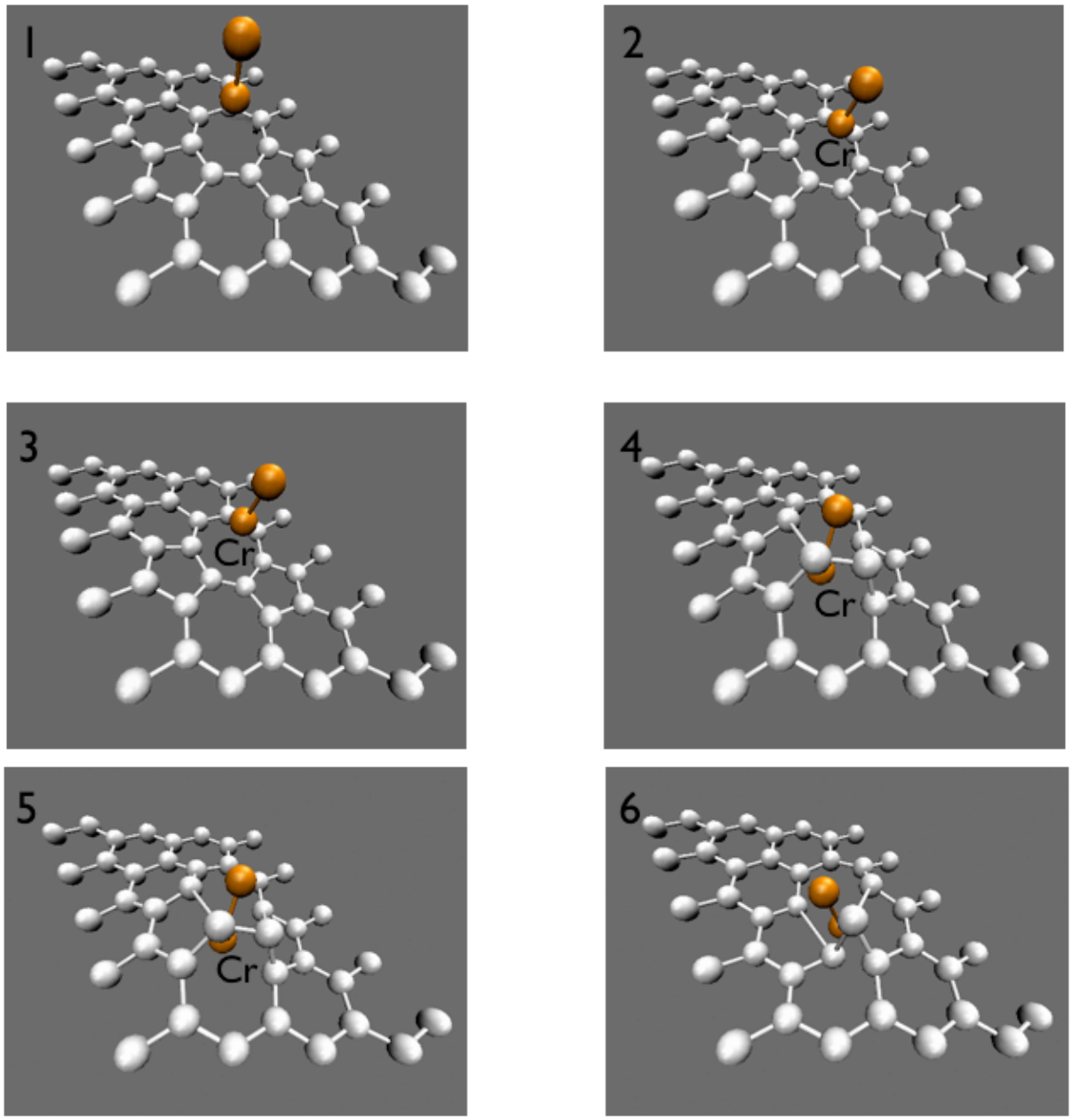


Figure 5.20: Optimized geometries in the configurations 1,2,3,4,5,6 in the case of SW defect (see Fig. 5.18). The z coordinate of the Cr atoms labeled in the figures was kept fixed during the optimizations.

5.6 Formation and dissociation of a Cr dimer: free energies

In section 2.7 we had seen that the free energy difference between two states ξ_N, ξ_1 can be calculated using the formula:

$$W(\xi_N) - W(\xi_1) = \int_{\xi_1}^{\xi_N} d\xi' \frac{\langle Z^{-1/2}[\lambda + k_B T G]\rangle_{\xi'}}{\langle Z^{-1/2}\rangle_{\xi'}} \tag{5.6}$$

In the cases reported in this section the constraint coordinate ξ is always represented has a distance between two atoms so that $Z = 1$ and $G = 0$ and Eq. 5.6 writes:

$$W(\xi_N) - W(\xi_1) = \int_{\xi_1}^{\xi_N} d\xi' \langle \lambda \rangle_{\xi'} \tag{5.7}$$

Obviously, to calculate this integral one has to divide the interval $[\xi_1, \xi_N]$ in subintervals so as to obtain the free energy differences as:

$$W(\xi_2) - W(\xi_1) \simeq \sum_{i=1}^{N} \frac{1}{2}\Big(\langle \lambda \rangle_{\xi_i} - \langle \lambda \rangle_{\xi_{i-1}}\Big)\Big(\xi_i - \xi_{i-1}\Big) = \sum_{i=1}^{N} \Delta_i d\xi_i \tag{5.8}$$

where $\{\xi_2, ..., \xi_n\}$ is a partition of the interval $[\xi_1, \xi_N]$ and:

$$\Delta_i = \frac{1}{2}\Big(\langle \lambda \rangle_{\xi_i} - \langle \lambda \rangle_{\xi_{i-1}}\Big)$$
$$d\xi_i = \xi_i - \xi_{i-1} \tag{5.9}$$

This approximation is better as it is smoother the behaviour of $\langle \lambda \rangle_{\xi_i}$ in function of ξ_i. Two free energy differences were calculated. The free energy difference of a dimer penetrating a graphene plane through a vacant site and the free energy difference of dimer formation when a Cr atom is placed at the center of an hexagon contiguous to a vacant site where another Cr atom is present. The coordination constraints for these reactions are graphically represented in Fig. 5.22. The case of dissociation reaction was studied using a two layer 4x4 graphite cell (63 atoms). The box length in the z direction is of about 10 Å and ξ represents the distance of the lower Cr atom and the C atom of the inferior graphite plane under the vacant site Fig 5.22. The case of formation reaction was studied with a single graphene plane (31 atoms) and the same length in z direction. In this case ξ represent the distance between Cr atoms. In previous sections the energy cut-off was 75 Ry. Since we are now

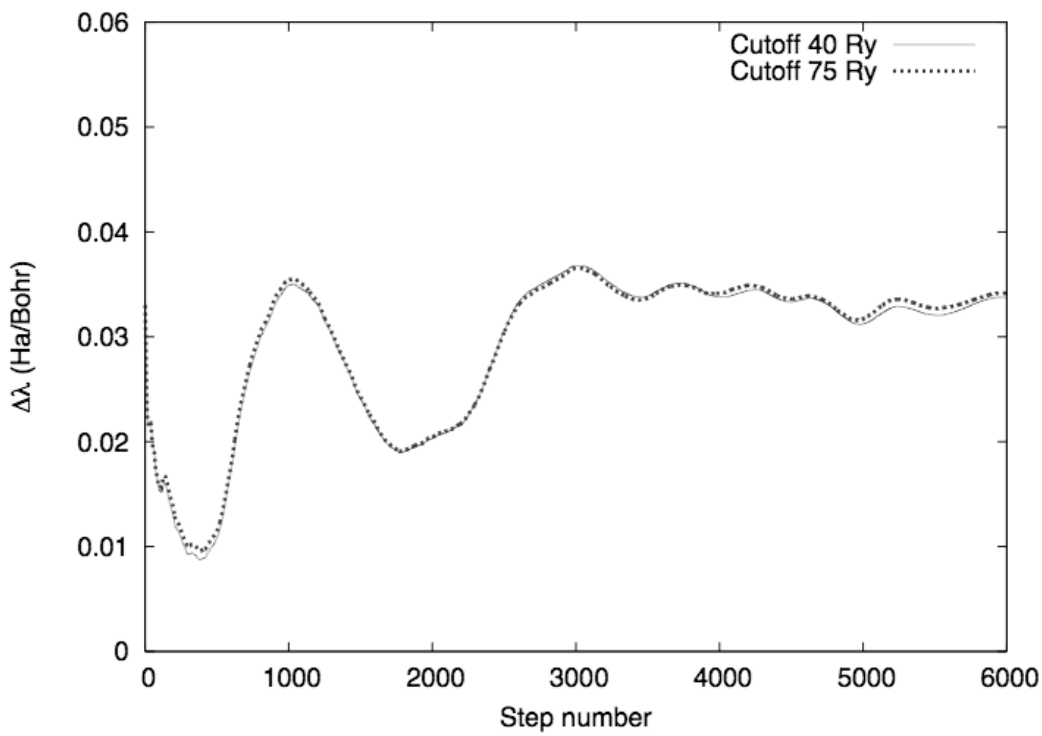


Figure 5.21: $2\Delta(9.367, 7.654) = \lambda(\xi = 9.367) - \lambda(\xi = 7.654)$ in function of simulated steps for different values of the energy cut-off.

interested in constraint force more than total energy differences, the cut-off convergence was tested with respect to this quantity. In particular what has to be correctly estimated in these MD runs are the differences in constraint forces between two consecutive values of constraint coordinate ξ Eq. 5.8. In Fig. 5.21 is reported the instantaneous difference of constraint force for two values of ξ is reported. From this figure it is clear that a cut-off of 40 Ry is sufficient.

CP molecular dynamics was used starting from the previous optimized geometry. The time step length was of about 0.1 fs. For every value of ξ an

MD run of about 1.7 ps was done. At first a MD run of 0.5 ps was done to equilibrate the system at 350 K using velocity rescaling. Then a run of about 1.2 ps was used to sample the constraint force λ thermalizing both the ionic and electronic system with a Nosé-Hoover chain thermostat. The ionic temperature was fixed at 350 K while target fictitious electronic kinetic energy was fixed through Eq. 3.56. Ionic and electronic temperature for a typical run are plotted in Fig. 5.23.

Results of calculations are reported in Tab 5.16 and Fig 5.24, 5.25. The correlation time and then the error on average constraint force were calculated using the technique reported in App. A.1. The statistical error on Δ_i is simply:

$$\sigma_{\Delta_i} = \frac{1}{2}\sqrt{\sigma^2_{\langle\lambda\rangle_{\xi_i}} + \sigma^2_{\langle\lambda\rangle_{\xi_{i-1}}}} \tag{5.10}$$

and statistical error on free energy difference between ξ_1 and $\xi_L \leqslant \xi_N$ is:

$$\sigma_{\Delta F(L,1)} = \sqrt{\sum_{i=1}^{L} \sigma^2_{\Delta_i} d\xi_i^2} \tag{5.11}$$

Both free energy differences are negative so the final states are thermodynamically stable. Free energy reaction barriers are of the order of some meV/Atom (Fig. 5.24, 5.25) which corresponds to some tens of Kelvin/Atom. The fact that the activation energy to form a Cr dimer on a graphene sheet is much smaller than the corresponding energy in vacuum has an easy interpretation: the energy needed to approach the electronic densities of the two Cr atoms is smaller because on graphene this electronic density is occupied to form bonds with graphene atoms.

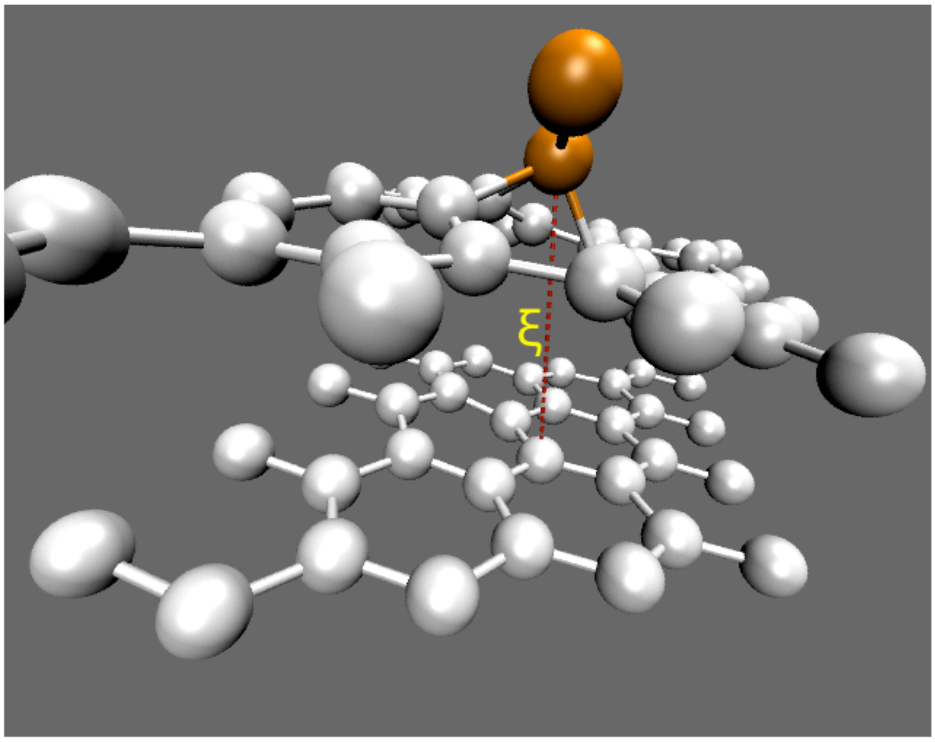
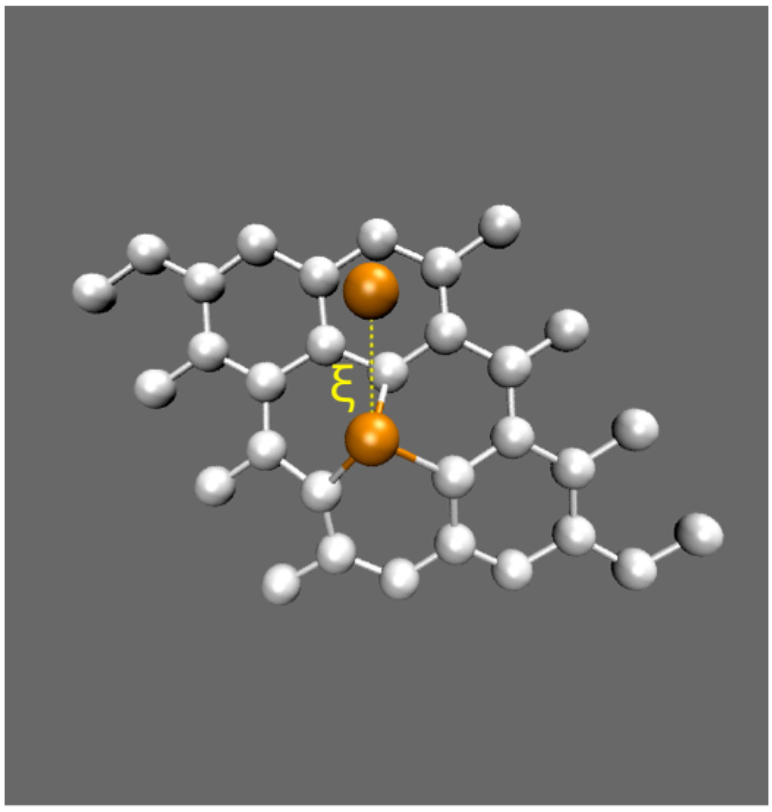


Figure 5.22: Graphical representation of the constraint coordinate in the two cases studied. Above, the constraint for the case of dissociation through a vacant site. Below, the constraint for the migration of Cr atom over a graphene plane.



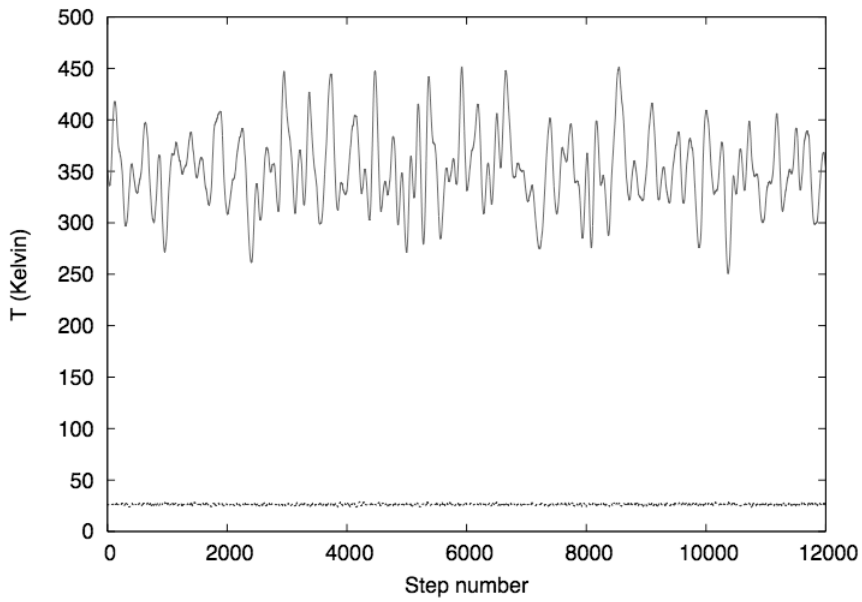
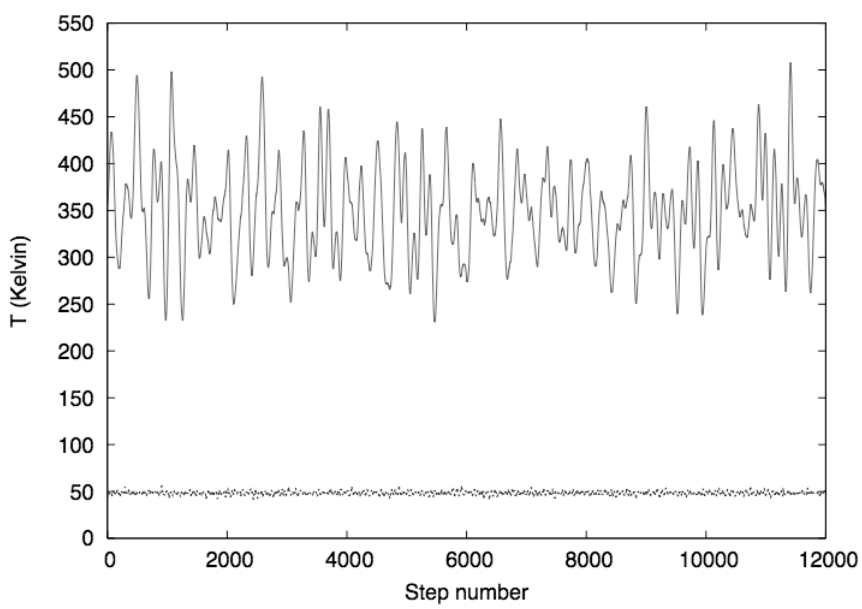


Figure 5.23: Typical behaviour of the ionic and fictitious electronic temperature. Above, the case of dissociation of the dimer, below, the case of formation. In red the ionic temperature.



ξ	$\langle\lambda\rangle$	σ^2_λ	t_c	σ_{t_c}	$\sigma^2_{\langle\lambda\rangle}$	σ_{Δ_i}	$\sigma_{\Delta_i} \times \Delta\xi_i$
9.367	9.89E-03	1.64E-04	187	7	2.5E-06	0.0	0.00
8.398	1.79E-02	1.84E-04	211	18	3.2E-06	1.2E-03	1.2E-03
7.654	3.80E-02	3.14E-04	529	6	1.4E-05	2.0E-03	7.0E-04
7.464	-1.42E-02	3.05E-04	378	6	9.6E-06	2.4E-03	4.6E-04
7.275	-1.23E-02	2.70E-04	228	11	5.1E-06	1.9E-03	3.6E-04
7.097	-1.11E-02	2.64E-04	227	5	5.0E-06	1.6E-03	2.8E-04
6.732	-6.00E-03	2.44E-04	223	5	4.5E-06	1.5E-03	5.5E-04
4.387	1.18E-03	1.80E-04	65	5	9.0E-07	1.1E-03	2.6E-03

Table 5.16: Results of the calculations for the cases studied. Atomic unit are used. The correlation time t_c is in number steps unit. Above, the case of dissociation, below, the case of formation.

ξ	$\langle\lambda\rangle$	σ^2_λ	t_c	σ_{t_c}	$\sigma^2_{\langle\lambda\rangle}$	σ_{Δ_i}	$\sigma_{\Delta_i} \times \Delta\xi_i$
5.0159	-2.59E-02	3.44E-04	452	16	1.3E-05	0.0	0.0
4.5408	-3.43E-02	2.89E-04	416	20	1.0E-05	2.4E-3	1.1E-3
3.9800	-2.77E-02	9.45E-05	203	7	1.6E-06	1.7E-3	9.5E-4
3.7667	-1.10E-03	9.50E-05	244	15	1.9E-06	9.3E-4	2.0E-4

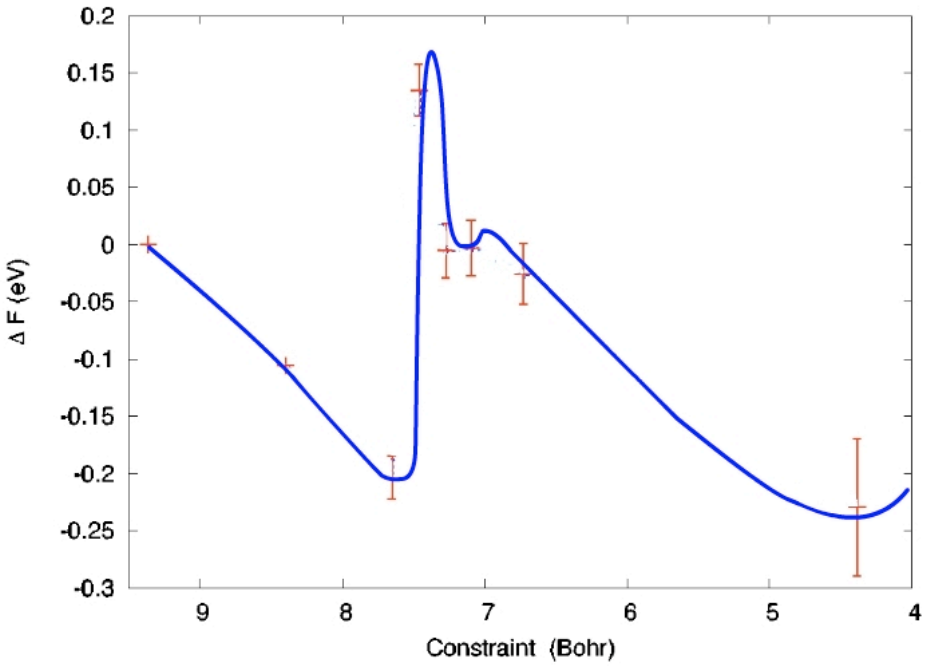
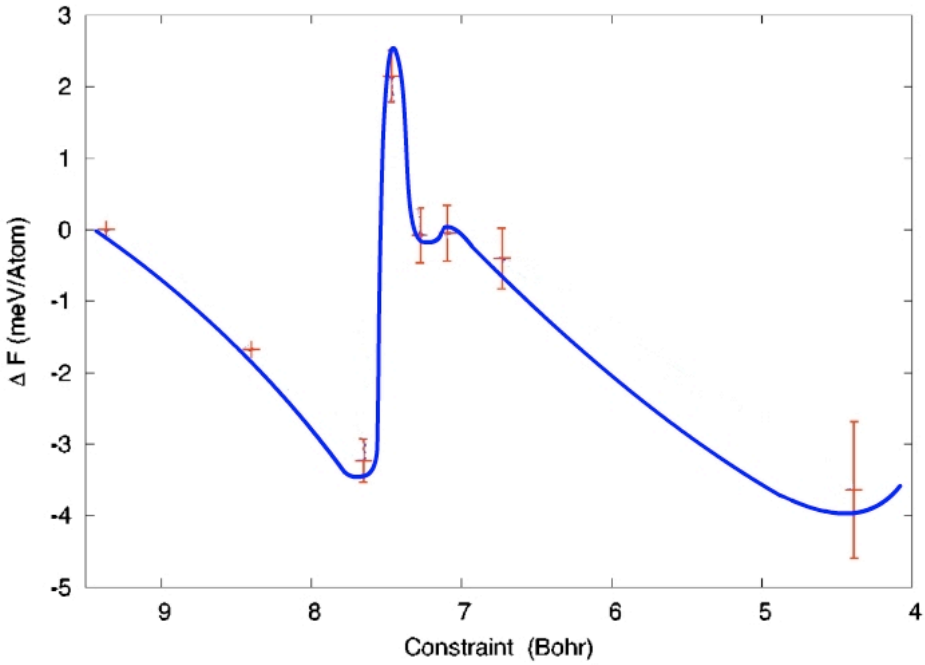


Figure 5.24: Free energy difference and statistical errors for the dissociation reaction in function of the reaction coordinate ξ (Fig. 5.22 above).



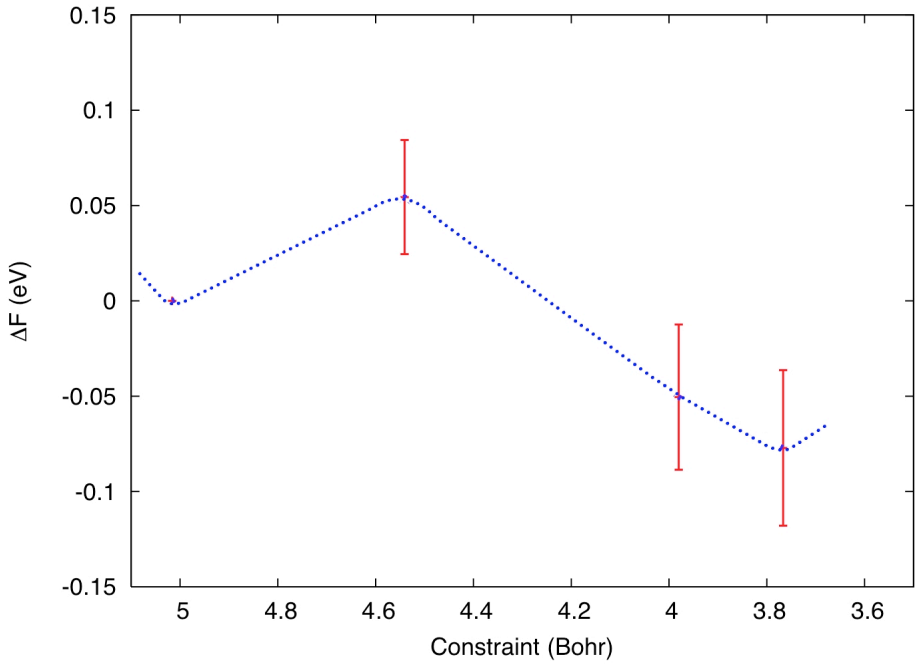
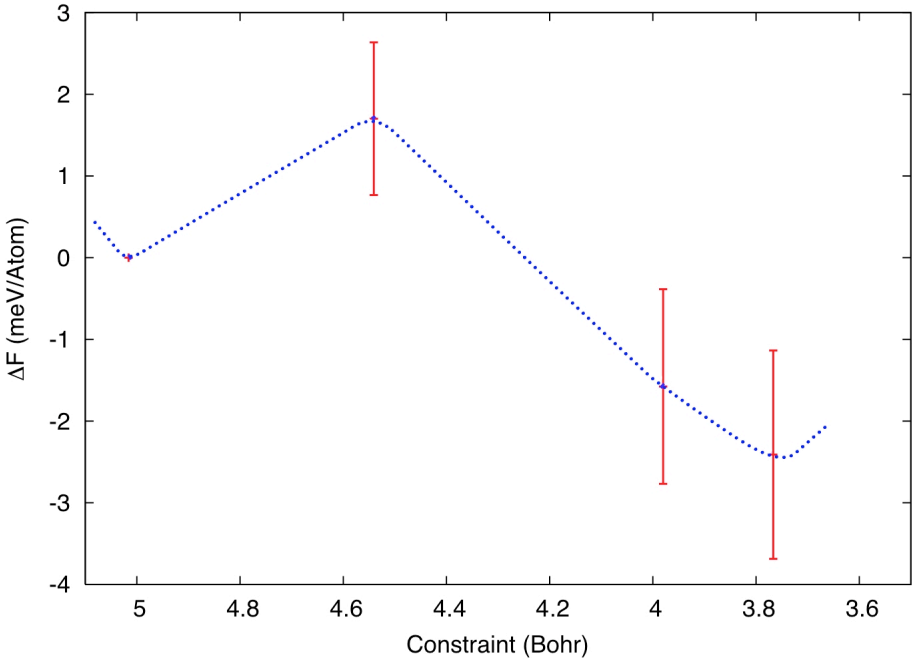


Figure 5.25: Free energy difference and statistical errors for the formation reaction in function of the reaction coordinate ξ (Fig. 5.22 below).



Some computational details

Although it was not emphasized elsewhere, ab-initio calculations require expensive machines to be performed in reasonable times. CASPUR's clusters were used in all calculations, except those with one or two atoms. In particular, I used Power-4, Power-5 and SX6 clusters. Power 4-5 clusters are composed by several processors, of the order of hundred, with a RISC (Reduced Instruction Set Computer) architecture. The SX6 cluster is composed by only eight processors with a vectorial architecture. A Power processor can execute only few operations in a clock cycle, while vectorial machines execute one operation for each RAM (Random Access Memory) address in a clock cycle. Obviously not all the codes can benefit of the vectorial architecture. However, even if each cluster is composed by several CPUs, one finds that, fixed the problem's size, the time needed to perform a calculation does not scale linearly with the number of CPUs used.

The calculations reported in Chapter 4 were done using the Abinit code which has a very detailed documentation and several utilities to investigate system's properties. This code scale linearly with the number of k-points used (in fact the calculations at each k-point are independent between themselves) so it works well for "small" systems which are in general studied with a large number of k-points. Abinit is not efficient for "large" systems when the calculations over the bands have to be efficiently parallelized. So I have used the CPMD code which is more performing for "large" systems. I used ultrasoft pseudopotentials whit CPMD instead of norm-conseving ones. The minimization of the total energy of the ionic system requires a number of geometry optimization steps which is not a priori predictable. Moreover, at each step of geometry optimization the electronic wavefunction (at fixed ion positions) must be calculated. This calculation requires a number of diagonalizations, or a number of direct minimization steps (depending on the algorithm used), which is not a priori predictable. In fortunate cases the number of these operations is of the order of 20,30 but it is not rare that one needs to do 100,150 steps. Each diagonalization or direct minimization requires a number of operations which grows like $N_{pw} \ln N_{pw}$ (N_{pw} represents the number of plane waves). The number of plane waves depend on the problem size in the following way:

$$N_{pw} \propto \Omega E_{cut}^{3/2} N_{band} \tag{5.12}$$

where Ω is the volume of the periodic box, N_{band} the number of bands (which

is proportional to the number of atoms and k-points) and E_{cut} the energy cut-off. This shows why one always tries to use a system, energy cut-off, etc... as small as possible.

Even performing CP dynamical simulations one cannot a priori estimate the time simulation length (t_s) needed to have a well converged average. That is in dynamical simulations which we performed we have to calculate (see Sec. 5.6):

$$\langle\lambda\rangle_\xi(t_s) = \frac{1}{t_s}\int_0^{t_s} \lambda(\tau)d\tau \tag{5.13}$$

for a set of values of the reaction coordinate ξ.

Let us now make some examples to have an idea about the computational cost. The calculation, performed with Abinit, of the Cr_3C_2 crystal which contains 20 atoms in its primitive cell had 8 k-points than the use of more than 8 CPUs (Abinit parallelization is not efficient over the bands) is of no utility. At first the atomic positions at fixed cell dimensions were optimized, this required 50 geometry optimization steps. Subsequently, both atomic positions and lattice parameters were optimized with other 35 optimization steps. The total calculation required 2 days with the Power4 cluster. However the convergence properties of the wavefunction in this situation were good (20-30 steps) since it is a bulk system and atoms are very close to their equilibrium configurations. Between the calculations performed with CPMD those reported in Fig. 5.18 required from 24 to 48 hours for each point in the fortunate cases (these were done in Γ point approximation and more than 4 CPUs, using the SX6 cluster, produced small performance improvement) and also a week in the unfortunate ones. The cost of each $\langle\lambda\rangle_\xi$ in the free energies calculations of Sec. 5.6 (starting from the optimized geometry at fixed z) in the case of the dissociation reaction, was of about 24 hours (16 Power5 CPUs).

Conclusions

The aim of this work was to have an understanding of the mechanical properties of MWCNTs under repeated twisting cycles. In particular we concentrated on the interaction between Chromium and graphene and on how this interaction can produce links between MWCNT's walls. This interaction seems to be crucial in characterizing the observed behaviour of MWCNTs (see the Chapter 1). Chromium in fact is evaporated on outer MWCNT's walls in all experiments to which we refer [71, 72, 49] because, as experimentally observed, it binds strongly to MWCNTs.

The calculations have confirmed that Cr atoms have strong interaction with graphene surface, both defective or not, and with two sheets of graphene if it is intercalated between them. It is then easy for Chromium to create links between MWCNT's walls. It is also clear that these simulations characterize well CNTs where the walls interaction is lower as their deformation properties indicate. In fact the geometry is easily distorted in z directions and a Cr atom can shift graphite planes. This simply shows as VdW forces are badly reproduced in DFT calculations.

The bond energies of Cr atom on graphene are all similar (0.4 eV of max difference). This suggests that once a Cr atom bounds to a plane, it can easily move on the surface. For Cr atoms intercalated through graphite planes, the max difference of bound energy is of about 1.3 eV, which is also not so high. The bound energies of intercalated Cr atoms are higher with respect to on surface energies. In all cases bond lengths are of the same order of magnitudes of typical Cr-C compounds, that is about 2 Å. These calculations were done with norm-conserving pseudopotentials and the Abinit code. The Cr pseudopotential was rebuilt since the library one has poorly reproduced some important physical quantities (Chapter 4).

Once the strong Cr-C interaction was clear, as also the calculations with ultrasoft pseudopotentials and the CPMD code was confirmed, we concentrated on the possible mechanisms which could generate the Chromium penetration inside MWCNT's walls. As suggested by calculations, a Cr atom cannot penetrate a graphene sheet even in presence of a vacancy. In fact even if one pushs a Cr atom in the center of a vacancy and locks its coordinates, graphene deforms to produce the same bonds obtained with a free Cr.

We have then tried with the smallest composite of Cr atoms: the Cr dimer. This was suggested when observing that a dimer not initially stretched in a

vacant site, dissociates Fig. 5.19. Some Cr dimers should also be present in the Cr gas evaporated onto MWCNTs. However while the activation energy required to form a dimer calculated in vacuum was of about 3 eV the calculated free energy barrier to form a dimer on a graphene surface was less than 0.1 eV. This suggests that dimers can form on graphene surface and can be interpreted in this way: the energy needed to approach the electronic densities of the two Cr atoms is smaller because on graphene this electronic density is occupied to form bonds with graphene atoms. The interaction of Cr dimers and graphene are also strong but in this case a defect on surface, in particular a vacancy, produces stronger bonds: on a perfect surface the higher binding energy was of -0.3 eV, on an SW defect -0.7 eV and on a vacant site -4.9 eV (Chapter 5). The preferred configuration for a dimer must then be on a vacant site. The activation energies of a dimer which, penetrating a defective graphene sheet, dissociates were calculated. The defects considered were a vacancy and a SW defect. The final state of a dimer stretched through a SW defect is not the most stable one and so it is not thermodynamically stable. In the case of vacancy the activation energy is 1.3 eV and the final states is the most stable one. The free energy difference of a dimer dissociating through a vacant site was calculated. It was found that $\Delta F \simeq -0.2$ eV and the free energy barrier was of about 0.35 eV.

There is a large number of investigations which could be done on these systems: at first, even if we found some processes which can explain the way in which Cr intercalates between graphene sheets, the reasons of the cascade mechanisms is still not clear. Secondly, since stable configurations of Cr atom intercalated between graphene or graphite sheets were found, it should be also interesting to understand how much this concentration can be increased to obtain stable composite and analyze their elastic properties. The electronic transport properties of CNTs or MWCTNs doped with Cr atoms or dimer should also be interesting. Finally the interface effects of Chromium crystal with CNTs should be investigated.

Let us now focus on possible applications of what reported in Sec. 2.6. The algorithm has to be better studied to understand if it effectively allows a larger time step. If this is the case, one can gain about 50% of CPU time. Another interesting application is the following: MC methods are faster to estimate static quantities. The reason is that in the MC sampling not correlated phase-space points are used to make average. Conversely in MD statistical errors are penalized by correlations. In other words, with the same number of phase-space points, one set obtained with a MD and another with MC sampling, the sta-

tistical error in the MC case is smaller. To treat constrained systems with MC techniques is very hard. For this reason (the total energy is a constraint) the (N, V, E) ensemble is never used with MC. An high increase of the time-step in the SHAKE-like algorithm can produce almost not correlated phase-space points. This can also be done considering higher order terms in the expression used to calculate λ. For the same reason (the presence of constraints) polymers are always treated in MD. The generalization of the SHAKE-like algorithm to also include additional holonomic constraints is straightforward. Using the algorithm in Nosé-Hoover equations can produce in the same way not correlated, canonical distributed, phase-space points, also in presence of additional holonomic constraints.

The technique exposed in Sec. 3.6 should also be applied for arbitrary temperatures. Let us consider the case of a one particle system. Instead of the Trotter factorization, one should write:

$$Tr[e^{-\beta H}] = \int dx \langle x|e^{-\frac{\beta}{n}H} \cdots e^{-\frac{\beta}{n}H}|x\rangle =$$

$$= \int dx dx_1 \cdots dx_n K_{\beta/n}(x - x_1) K_{\beta/n}(x_1 - x_2) \cdots K_{\beta/n}(x_N - x) \tag{5.14}$$

where:

$$K_{\beta/n}(x_{N-1} - x_N) = \langle x_{N-1}|e^{-\frac{\beta}{n}H}|x_N\rangle \tag{5.15}$$

If n is large enough, this kernel can be evaluated at hight temperature so that:

$$\langle x_{N-1}|e^{-\frac{\beta}{n}H}|x_N\rangle \propto e^{-\beta/n V_{eff}(x_{N-1} - x_N)} \tag{5.16}$$

Then Eq. 5.14 represents a closed polymer composed by n particles interacting through V_{eff}. This partition function can be computed by MD.

Acknowledgments

I want to thank F. Cleri who has introduced me to ab-initio calculations and with A. Di Carlo gave me useful suggestions. I am very grateful to G. Ciccotti who discussed with me what is reported in section 2.6, 3.6. I want also to thank CASPUR for financial support and some members of CASPUR HPC staff who have helped me in technical problems: A. Federico, L. Ferraro and M. Rosati. In particular I want to thank S. Meloni for some discussions regarding ab-initio simulations.

Appendices

A.1 Statistical errors in time averages

In section 5.6 we needed to calculate the statistical error of average constraint force obtained with a MD run. Since the values $\lambda(t_k)$ are not statistically independent, indicating with n_r the number of sampled values of λ during the MD run ($n_r * dt = t_r$, where dt is time-step length), we cannot use the well known formula:

$$\sigma^2(\langle\lambda_{run}\rangle) = \frac{1}{n_r}(\langle\lambda^2_{run}\rangle - \langle\lambda_{run}\rangle^2) \tag{A.1}$$

where:

$$\langle\lambda\rangle_{run} = \frac{1}{t_r}\sum_{i=1}^{r}\lambda(t_i) \tag{A.2}$$

let us write σ^2 in the following way:

$$\begin{aligned}\sigma^2(\langle\lambda_{run}\rangle) &= \frac{1}{t_r^2}\int_0^{t_r}\int_0^{t_r} dtdt'\langle[\lambda(t)-\langle\lambda\rangle][\lambda(t')-\langle\lambda\rangle]\rangle = \\ &= \frac{1}{t_r^2}\int_0^{t_r}\int_0^{t_r} dtdt' C_\lambda(t-t')\end{aligned} \tag{A.3}$$

where C_λ is the time correlation function of λ (what follows can be found in [2, 20] and reference therein). Define the correlation time t^c_λ as:

$$t^c_\lambda = \frac{1}{2}\int_{-\infty}^{\infty}\frac{C_\lambda(t)}{C_\lambda(0)}dt \tag{A.4}$$

Supposing in A.3 that $t_r \gg t^c_\lambda$ we obtain:

$$\begin{aligned}\sigma^2(\langle\lambda_{run}\rangle) \approx \frac{1}{t_r}\int_0^{\infty} C_\lambda(t)dt = \frac{1}{2t_r}\int_{-\infty}^{\infty} C_\lambda(t)dt = \\ = \frac{t^c_\lambda}{t_r}C_\lambda(0) = \frac{t^c_\lambda}{t_r}(\langle\lambda^2_{run}\rangle - \langle\lambda_{run}\rangle^2)\end{aligned} \tag{A.5}$$

It follows that from the knowledge of t_λ^C an estimate of the variance of $\langle\lambda_{run}\rangle$ can be obtained. Suppose now to divide the sampled time t_r in n_b block of length t_b ($n_b * t_b = t_r$). For each block we define:

$$\langle\lambda\rangle_b = \frac{1}{t_b}\sum_{i=1}^{t_b}\lambda(t_i) \tag{A.6}$$

Then the variance of $\langle\lambda\rangle_b$ writes:

$$\sigma^2(\langle\lambda\rangle_b) = \frac{1}{n_b}\sum_{b=1}^{n_b}\Big(\langle\lambda\rangle_b - \langle\lambda\rangle_{run}\Big)^2 \tag{A.7}$$

Increasing the block length, $\langle\lambda\rangle_b$ becomes uncorrelated, then from Eq. A.3 we expect that:

$$\sigma^2(\langle\lambda\rangle_b) \propto \frac{1}{t_b} \tag{A.8}$$

An estimate of t_λ^C can be obtained as:

$$t_\lambda^C = \lim_{t_b\to\infty} s(t_b) = \lim_{t_b\to\infty} t_b\,\frac{\sigma^2(\langle\lambda\rangle_b)}{\sigma^2(\langle\lambda\rangle_{run})} \tag{A.9}$$

The function $s(t_b)$ is called statistical efficiency. From Eq. A.3 we obtain:

$$\sigma^2(\langle\lambda_{run}\rangle) = \frac{1}{t_b}\cdot(\langle\lambda_{run}^2\rangle - \langle\lambda_{run}\rangle^2)\cdot\lim_{t_b\to\infty} s(t_b) \tag{A.10}$$

To estimate practically t_c we had calculated $s(n_b)$ for different values of n_b. When the plateau was reached we took four values of $s(n_b)$ with n_b increased of fifty steps for each point. Then the asymptotic value of $s(n_b)$, was obtained fitting the asymptote with four points. An example of the behaviour of $s(n_b)$ is showed in Fig. A.1. The estimated t_c in number of steps unit is also reported in the figure.

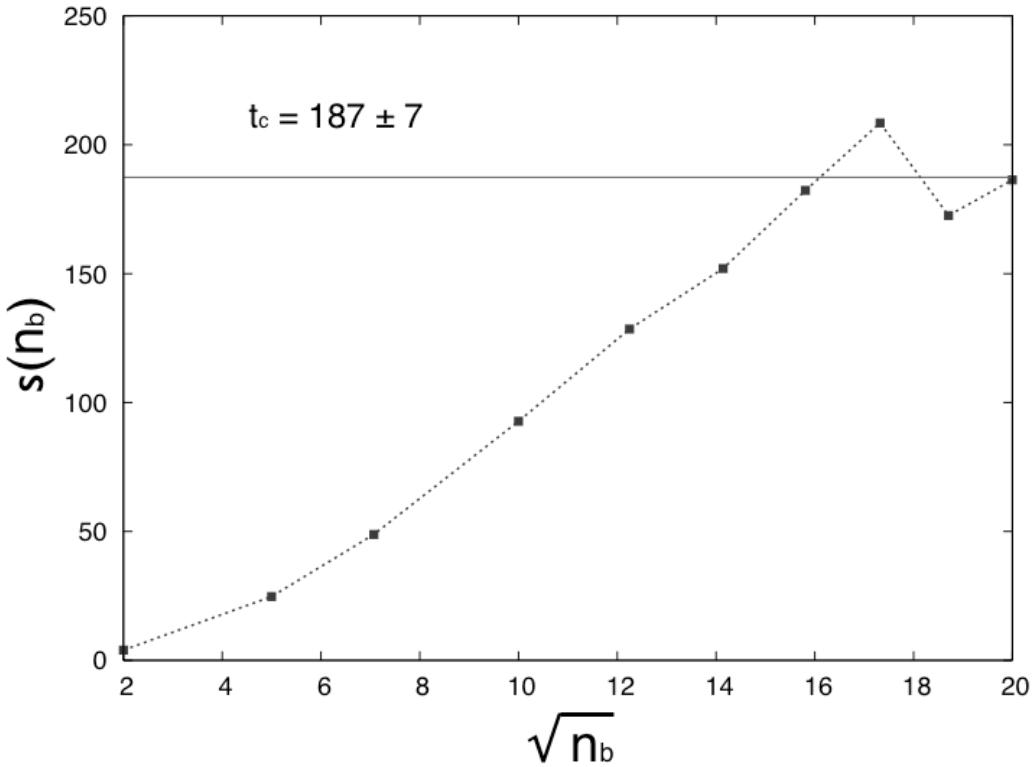


Figure A.1: An example of t_c estimation measured in number of steps. The x-axis represent the square root of points per block.

A.2 Distortion and divergence

To measure the variation of atomic position in a configuration $\mathbf{r'_i}$ relatively to a reference configuration $\mathbf{r_i}$ we introduce the distortion as:

$$D_x = \frac{1}{N}\sum_{i=1}^{N}(x'_i - x_i)^2$$
$$D_y = \frac{1}{N}\sum_{i=1}^{N}(y'_i - y_i)^2$$
$$D_z = \frac{1}{N}\sum_{i=1}^{N}(z'_i - z_i)^2 \tag{A.11}$$

where N is the number of atoms. Analogously we define the divergence of two trajectories of N_{step} configurations $\mathbf{r'^k_i}, \mathbf{r^k_i}$, with $k = 1, ..., N_{step}$ as:

$$D = \sum_{k=1}^{N_{step}}\sum_{i=1}^{N}|r'^k_i - r^k_i|^2 \tag{A.12}$$

The average separation of two trajectories of length N_{step} is:

$$D_{avg} = D/N_{step} \tag{A.13}$$

A.3 Notations

a, b, c	Lattice parameters
B	Bulk modulus
BFGS	Broyden-Fletcher-Goldfarb-Shanno
BO	Born-Oppenheimer
CNT	Carbon nanotubes
CP	Car-Parrinello
E_{ch}	Cohesive energy
eV	Electronvolt
ΔEs	Variation of electrostatic energy
ΔE_{GGA}	Variation of gradient corrected energy
ΔLPs	Variation of local pseudopotential energy
$\Delta NLPs$	Variation of non-local psudopotential energy
ΔXC	Variation of exchange correlation energy
D	Trajectories divergence
D_x, D_y, D_z	Distortion/atoms in the three spatial directions
GGA	Generalized gradient approximation
GPa	Giga Pascal
H	Hamman
Ha	Hartree
k_B	Boltzmann constant
KB	Kleinman-Bylander
KS	Kohn-Sham
LDA	Local density approximation
MC	Monte Carlo
MD	Molecular dynamics
MWCNT	Multi-walled carbon nanotubes
NLCC	Nonlinear core correction
PBE	Perdew-Burke-Ernzerhof
PW92	Perdew-Wang 1992
r_l	Matching distance for l pseudopotential channel
r_{nlc}	Matching distance for core density in NLCC
Ry	Rydberg
σ	Standard deviation
σ_{ik}	i, k component of the stress tensor

SM	Statistical mechanics
T	Temperature
TM	Troullier-Martins
V	Volume
XC	Exchange Correlation

References

[1] M. J. Allen and D. J. Tozera: *Helium dimer dispersion forces and correlation potentials in density functional theory*, J. Chem. Phys. **117, 24**, 11113 (2002).

[2] M. P. Allen and D. J. Tildesley: Computer Simulation of Liquids (Calarendon Press) (1987).

[3] H. C. Andersen: *Rattle: a 'velocity' version of the SHAKE algorithm for molecular dynamics calculations*, J. Comp. Phys. **52, 1**, 24 (1983).

[4] G. B. Bachelet, D. R. Hamann and M. Schüter: *Pseudopotentials that work: From H to Pu*, Phys. Rev. B **26, 8**, 4199 (1982).

[5] G. Benetin and A. Giorgilli: *On the Hamiltonian interpolation of near-to-the-identity symplectic mappings with application to symplectic integration aigorithms*, J. Stat. Phys. **74, 5-6**, 1117 (1994).

[6] J. Bernholc and N. A. W. Holzwarth: *Local spin-density description of multiple metal-metal bonding: Mo_2 and Cr_2*, Phys. Rev. Lett. **50**, 1451 (1983).

[7] P. E. Bl0̈chl and M. Parrinello: *Adiabaticity in first principles molecular dynamics*, Phys. Rev. B **45, 16**, 9413 (1992).

[8] B. H. Bransden and C. J. Joachain: Physics of Atoms and Molecules (Addison Wesley Longman Limited) (1983).

[9] R. Car and M. Parrinello: *Unified approach for molecular dynamics and density-functional theory*, PRL **55, 22**, 2471 (1985).

[10] G. Ciccotti and M. Ferrario: *Rare events by constrained molecular dynamics*, J. Mol. Liq. **89, 1**, 1 (2000).

[11] G. Ciccotti and M. Ferrario: *Blue Moon approach to rare events*, Mol. Simul. **30, 11-12**, 787 (2004).

[12] I. P. Cornfeld, S. V. Fomin and Y. G. Sinai: Ergodic Theory (Springer-Verlag) (1980).

[13] H. Dachsel, R. J. Harrison and D. A. Dixon: *Multireference configuration interaction calculations on Cr_2 passing the one billion limit in MRCI/MRACPF calculations*, J. Phys. Chem. A **103**, 152 (1999).

[14] A. Di Carlo, M. Monteferrante, P. Podio-Guidugli, V. Sansalone and L. Teresi: *How (and why) twisting cycles make individual MWCNTs stiffer*, Molecular Nanostructures IWEPNM 2004 (AIP Conf. Proc.), volume 685, 577 (H. Kuzmany et al. and eds.) (2004).

[15] D. J. Evans, W. G. Hoover, B. H. Failor, B. Moran and A. J. C. Ladd: *Nonequilibrium molecular dynamics via Gauss's principle of least constraint*, Phys. Rev. A **28, 2**, 1016 (1983).

[16] G. S. Ezra: *On the statistical mechanics of non-Hamiltonian systems: the generalized Liouville equation, entropy, and time-dependent metrics*, J. Math. Chem. **35, 1**, 29 (2004).

[17] Y. Fan, R. Kalaba, H. Natsuyama and F. Udwadia: *Reflections on the 175th anniversary of Gausss principle of least constraint*, J. Optim. Theory Appl. **127, 3**, 475 (2005).

[18] A. Fasano and S. Marmi: Meccanica Analitica (Bollati Boringhieri) (1994).

[19] A. Fennimore, T. Yuzvinsky, W.-Q. Han, M. Fuhrer, J. Cumings and A. Zettl: *Rotational actuators based on carbon nanotubes*, Nature **424**, 408 (2003).

[20] D. Frenkel and B. Smit: Understanding molecular simulations (Academic Press) (1996).

[21] M. Fuchs and M. Scheffler: *Ab initio pseudopotentials for electronic structure calculations of poly-atomic systems using density-functional theory*, Comput. Phys. Comm. **119, 1**, 67 (1999).

[22] H. Goldstein, C. Poole and J. Safko: Classical Mechanics (John Wiley and Sons) (1965).

[23] X. Gonze, P. Käckell and M. Scheffler: *Ghost states for separable, norm-conserving, ab initio pseudopotentials*, Phys. Rev. B **41, 17**, 12264 (1990).

[24] X. Gonze, R. Stumpf and M. Scheffler: *Analysis of separable potentials*, Phys. Rev. B **44, 16**, 8503 (1991).

[25] S. K. Gray, D. W. Nold and B. G. Sumpter: *Symplectic integrators for large scale molecular dynamics simulations: A comparison of several explicit methods*, J. Chem. Phys. **101, 5**, 4062 (1994).

[26] E. Hairer and C. Lubich: *The life-span of backward error analysis for numerical integrators*, Numer. Math. **76, 4**, 441 (1997).

[27] P. Hohenberg and W. Khon: *Inhomogeneous Electron Gas*, Phys. Rev. **136, B3**, 864 (1964).

[28] W. G. Hoover: *Canonical dynamics: Equilibrium phase-space distributions*, Phys. Rev. A **31, 3**, 1695 (1985).

[29] W. G. Hoover: *Constant pressure equations of motion*, Phys. Rev. A **34, 3**, 2499 (1986).

[30] K. Huang: Statistical Mechanics (John Wiley and Sons) (1963).

[31] M. Huhtala, A. V. Krasheninnikov, J. Aittoniemi, S. J. Stuart, K. Nordlund and K. Kaski: *Improved mechanical load transfer between shells of multi-walled carbon nanotubes*, Physical Review B **70**, 45404 (2004).

[32] G. P. Kerker: *Non-singular atomic pseudopotentials for solid state applications*, Solid St. Phys. **13, 9**, 189 (1980).

[33] A. Kis, G. Csnyi, J.-P. Salvetat, T.-N. Lee, E. Couteau, A. J. Kulik, W. Benoit, J. Brugger and L. Forrò : *Reinforcement of single-walled carbon nanotube bundles by intertube bridging*, Nature Mater. **3**, 153 (2004).

[34] W. Kohn and L. J. Sham: *Self-consistent equations including exchange and correlation effects*, Phys. Rev. **140, 4A**, 1133 (1965).

[35] K. Laasonen, R. Car, C. Lee and D. Vanderbilt: *Implementation of ultrasoft pseudopotentials in ab initio molecular dynamics*, Phys. Rev. B **43, 8**, 6796 (1991).

[36] K. Laasonen, A. Pasquarello, R. Car, C. Lee and D. Vanderbilt: *Car-Parrinello molecular dynamics with Vanderbilt ultrasoft pseudopotentials*, Phys. Rev. B **47, 16**, 10142 (1993).

[37] B. J. Leimkuhler and R. D. Skeel: *Symplectic numerical integrators in constrained Hamiltonian systems*, J. Comp. Phys. **112, 1**, 117 (1994).

[38] S. Letardi, M. Celino, F. Cleri and V. Rosato: *Atomic hydrogen adsorption on a Stone-Wales defect in graphite*, Surf. Sci. **496, 1-2**, 33 (2002).

[39] J. Lu: *Elastic properties of carbon nanotubes and nanoropes*, Phys. Rev. Lett. **79**, 1297 (1997).

[40] Y. Ma, P. O. Lehtinen, A. S. Foster and R. M. Nieminen: *Magnetic properties of vacancies in graphene and single-walled carbon nanotubes*, New J. Phys. **6, 68**, 1 (2004).

[41] H. J. Monkhorst and J. D. Pack: *Special points for Brillouin-zone integrations*, Phys. Rev. B **13, 12**, 13 (1976).

[42] M. B. Nardelli: *Electronic transport in extended systems: Application to carbon nanotubes*, Physical Review B **60, 11**, 7828 (1999).

[43] M. B. Nardelli and J. Bernholc: *Mechanical deformations and coherent transport in carbon nanotubes*, Physical Review B **60, 24**, 16338 (1999).

[44] M. B. Nardelli, B. Yakobson and J. Bernholc: *Brittle and ductile behavior in carbon nanotubes*, Phys. Rev. Lett. **81**, 4656 (1998).

[45] M. B. Nardelli, B. I. Yakobson and J. Bernholc: *Mechanism of strain release in carbon nanotubes*, Phys. Rev. B **57**, 4277 (1998).

[46] S. Nosé : *A molecular dynamics method for simulation in the canonical ensemble*, Mol. Phys. **52, 2**, 255 (1984).

[47] P. Olsson, I. Abrikosov, L. Vitos and J. Wallenius: *Ab initio formation energies of Fe-Cr alloys*, J. Nucl. Mater. **321**, 84 (2003).

[48] S. Papadakis, P. Williams, M. Falvo, R. Superfine and S. Washburn: *Mechanics of nanotubes and nanotube-based devices*, Molecular Nanostructures (AIP Conf. Proc.), volume 685, 577 (H. Kuzmany et al. and eds.) (2003).

[49] S. J. Papadakis, A. R. Hall, P. A. Williams, L. Vicci, M. R. Falvo, R. Superfine and S. Washburn: *Resonant oscillators with carbon-nanotube torsion springs*, Phys. Rev. Lett. **93**, 146101 (2004).

[50] G. Pastore, G. Smargiassi and F. Buda: *Theory of ab initio molecular-dynamics calculations*, Phys. Rev. A **44, 10**, 6334 (1991).

[51] M. C. Payne, M. P. Teter, D. C. Allan, T. A. Arias and J. D. Joannopoulos: *Iterative minimization techniques for ab-initio total energy calculations: molecular dynamics and conjugate gradients*, Rev. Mod. Phys. **64, 4**, 1045 (1992).

[52] M. C. Payne, M. P. Teter, D. C. Allan, T. A. Arias and J. D. Joannopoulos: *Iterative minimization techniques for ab-inition total energy calculations: molecular dynamics and conjugate gradients*, Rev. Mod. Phys. **64, 4**, 1045 (1992).

[53] S. Reich: *Backward error analysis for numerical integrators*, SIAM J. Numer. Anal. **36, 5**, 475 (1999).

[54] L. Rosso, P. Minà ry, Z. Zhu and M. E. Tuckerman: *On the use of the adiabatic molecular dynamics technique in the calculation of free energy profiles*, J. Chem. Phys. **116, 11**, 4389 (2002).

[55] J. P. Ryckaert, G. Ciccotti and H. J. C. Berendsen: *Numerical integration of the Cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes*, J. Comp. Phys. **23, 3**, 327 (1977).

[56] D. J. Singh and B. M. Klein: *Electronic structure, lattice stability, and superconductivity of CrC*, Phys. Rev. B **46**, 14969 (1992).

[57] M. Sprik and G. Ciccotti: *Free energy from constrained molecular dynamics*, J. Chem. Phys. **109, 18**, 7737 (1998).

[58] I. Stich, R. Car, M. Parrinello and S. Baroni: *Conjugate gradient minimization of the energy functional: a new method for electronic structure calculations*, Phys. Rev. B **39, 8**, 4997 (1989).

[59] P. Tangney and S. Scandolo: *How well do Car-Parrinello simulations reproduce the Born-Oppenheimer surface? theory and examples*, J. Chem. Phys. **116, 1**, 14 (2002).

[60] R. Telling, C. Ewels, A. El-Barbary and M. Heggie: *Wigner defects bridge the graphite gap*, Nature Mater. **2**, 333 (2003).

[61] N. Troullier and J. L. Martins: *Efficient pseudopotentials for plane-wave calculations*, Phys. Rev. B **43, 3**, 1993 (1991).

[62] M. Tuckerman, B. J. Berne and G. J. Martyna: *Reversible multiple time scale molecular dynamics*, J. Chem. Phys. **97, 3**, 1990 (1992).

[63] M. E. Tuckerman: *Path integration via molecular dynamimcs*, Quntum simulations of complex many-body systems: from theory to applications **10**, 269 (2002).

[64] M. E. Tuckerman, Y. Liu, G. Ciccotti and G. J. Martyna: *Non-Hamiltonian molecular dynamics: Generalizing Hamiltonian phase space principles to non-Hamiltonian systems*, J. Chem. Phys. **115, 4**, 1678 (2001).

[65] M. E. Tuckerman and M. Parrinello: *Integrating the Car-Parrinello equations. i. basic integration techniques*, J. Chem. Phys. **101, 2**, 1302 (1994).

[66] M. E. Tuckerman and M. Parrinello: *Integrating the Car-Parrinello equations. ii. multiple time scale techniques*, J. Chem. Phys. **101, 2**, 1316 (1994).

[67] D. Vanderbilt: *Soft self-consistent pseudopotentials in a generalized eigenvalue formalism*, Phys. Rev. B **41, 11**, 7892 (1990).

[68] M. Verstraete and X. Gonze: *Smearing scheme for finite-temperature electronic-structure calculations*, Phys. Rev. B **65, 35111** (2002).

[69] M. Watanabe and W. P. Reinhardt: *Direct dynamical calculation of entropy and free energy by adiabatic switching*, Phys. Rev. Lett. **65, 26**, 3301 (1990).

[70] E. Wigner: *On the quantum correction for thermodynamic equilibrium*, Physical Review **40**, 749 (1932).

[71] P. Williams, S. Papadakis, A. Patel, M. Falvo, S. Washburn and R. Superfine: *Torsional response and stiffering of individual multiwalled carbon nanotubes*, Phys. Rev. Lett. **89**, 255502 (2002).

[72] P. Williams, S. Papadakis, A. Patel, M. Falvo, S. Washburn and R. Superfine: *Fabrication of nanometer-scale paddle oscillators incorporating individual multiwalled carbon nanotubes*, Appl. Phys. Lett. **82**, 805 (2003).

[73] Q. Zhao, M. B. Nardelli and J. Bernholc: *Ultimate strength of carbon nanotubes: A theoretical study*, Physical Review B **65**, 144105 (2002).

[74] Q. Zhao, M. B. Nardelli, W. Lu and J. Bernholc: *Carbon nanotube-metal cluster composites: a new road to chemical sensors?*, Nano Lett. **5**, 847 (2005).

[75] U. Zimmerli, M. Parrinello and P. Koumoutsakos: *Dispersion corrections to density functionals for water aromatic interactions*, J. Chem. Phys. **120, 6**, 2693 (2004).