ON DENSITY FUNCTIONAL PERTURBATION THEORY WITH THE FINITE ELEMENT METHOD

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF ENGINEERING AND SCIENCE
OF BILKENT UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR
THE DEGREE OF
MASTER OF SCIENCE
IN
MECHANICAL ENGINEERING

By
Yunus Altıntop
August 2025

ON DENSITY FUNCTIONAL PERTURBATION THEORY WITH THE FINITE ELEMENT METHOD

By Yunus Altıntop

August 2025

We certify that we have read this thesis and that in our opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

İlker Temizer(Advisor)

Seymur Jahangirov

Osman Barış Malcıoğlu

Approved for the Graduate School of Engineering and Science:

Orhan Arıkan
Director of the Graduate School

ABSTRACT

ON DENSITY FUNCTIONAL PERTURBATION THEORY WITH THE FINITE ELEMENT METHOD

Yunus Altıntop
M.S. in Mechanical Engineering
Advisor: İlker Temizer
August 2025

Density Functional Perturbation Theory (DFPT) introduces a systematic framework to compute higher-order material properties based on Density Functional Theory (DFT) and linear response theory. However, most conventional implementations are intrinsically formulated for periodic structures, which may obscure flexibility of complex systems. In order to overcome the limitations, DFPT is reformulated for the first time within Finite Element Method (FEM), a real-space method. The models are developed for both all-electron and pseudopotential frameworks with the GGA and LDA exchange-correlation functionals. For aperiodic system implementations, benchmark calculations achieved chemical accuracy.

Keywords: Kohn-Sham Density Functional Theory, pseudopotentials, all-electron, exchange-correlation functional, real-space methods, interatomic force constant tensor, Sternheimer equation, gauge freedom, projector.

ÖZET

SONLU ELEMANLAR YÖNTEMİYLE YOĞUNLUK FONKSİYONELİ PERTÜRBASYON TEORİSİ ÜZERİNE

Yunus Altıntop
Makine Mühendisliği, Yüksek Lisans
Tez Danışmanı: İlker Temizer
Ağustos 2017

Yoğunluk Fonksiyoneli Pertürbasyon Teorisi (DFPT), Yoğunluk Fonksiyoneli Teorisi (DFT) ve doğrusal tepki teorisine dayalı olarak daha yüksek mertebeden malzeme özelliklerini hesaplamak için sistematik bir çerçeve sunar. Ancak, standart uygulamaları doğası gereği periyodik yapılar öngörmekte ve bu durum karmaşık sistemler için esnekliği sınırlayabilmektedir. Bu kısıtlamaların aşılması amacıyla, DFPT gerçek uzay yöntemi olan Sonlu Elemanlar Yöntemiyle (FEM) ilk kez formüle edilmiştir. Modeller, tüm elektron (all-electron) ve yalancı potansiyel (pseudopotential) yaklaşımları için, GGA ve LDA değiş-tokuş/ilinti (exchange-correlation) fonksiyonelleriyle geliştirilmiştir. Periyodik olmayan sistemler üzerindeki uygulamalarda, karşılaştırmalı hesaplamalar kimyasal doğruluğa ulaşmıştır.

Anahtar sözcükler: Kohn-Sham Yoğunluk Fonksiyoneli Teorisi, yalancı potansiyeller, tüm-elektron, değiş-tokuş/ilinti fonksiyoneli, gerçek-uzay yöntemleri, atomlar arası kuvvet sabiti tensörü, Sternheimer denklemi, ölçüt serbestliği, projektör.

Acknowledgement

I would like to express my deepest gratitude to my advisor, Prof. Dr. İlker Temizer, for his invaluable guidance, patience, and continuous support throughout my academic journey. His insight, encouragement, and dedication to research have been a constant source of inspiration, shaping both my scientific approach and personal growth.

I am profoundly thankful to my family for their unconditional love, understanding, and sacrifices that made this work possible. Their unwavering belief in me has been my greatest source of motivation.

I also wish to extend my heartfelt appreciation to my friends, whose encouragement, laughter, and companionship have lightened the challenges of this journey and made it an unforgettable experience.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Density Functional Theory (DFT) has emerged as a feasible alternative to the many-body Schrödinger equation for electronic structure calculations of atoms, molecules and crystals [1]. Over years, it has been improved by introducing system-specific pseudopotentials and more accurate exchange-correlation functionals [2, 3, 4]. Furthermore, it has been implemented in various numerical frameworks such as the *Gaussian basis* in real space and the *plane-wave* in reciprocal space [5, 6, 7].

Despite the variational structure of DFT, the Gaussian basis formulation lacks monotonic energy convergence upon basis refinement on its incomplete basis set [8]. In contrast, the plane-wave formulation achieves systematic convergence by increasing the *kinetic energy cutoff*. Even though the Gaussian basis is feasible for molecular systems, it is not suitable for periodic structures like crystals. On the other hand, the plane-wave basis intrinsically possess periodic boundary conditions making it natural tool for crystal structures. However, its molecular models require *supercells*, which are large unit cells, to isolate them from the contagious unit cell images.

Among alternative numerical methods in real space, *Finite Difference* (FD) and *Finite Element Method* (FEM) have been verified within the Kohn-Sham

DFT framework [9, 10]. Especially, it has been proven that FEM could achieve optimal energy convergence rates with respect to the *number of degree of freedom* (DOF). Moreover, the scope of DFT implementations within FEM framework has reached to energy, force, stress calculations in both aperiodic and periodic systems, and beyond [11, 12, 13]. However, *Density Functional Perturbation Theory* (DFPT) within FEM framework has not been implemented so far while it was studied within FD framework [14]. This study aims to implement basics of DFPT in aperiodic structures within FEM framework.

DFPT allows access to myriads of material properties considering linear response theory [15]. It variationally searches for the Kohn-Sham wavefunction perturbations minimizing even-order energy perturbations [16]. In this study, *interatomic force constant* (IFC) tensor, corresponding to the second-order energy perturbation with respect to ionic positions, is computed. It is useful in order to find vibrational modes of molecules, which allow to observe infrared and Raman spectra.

This thesis consists of six chapters. Chapter 1 provides a general picture of the study. Chapter 2 starts with a background on the development of DFT, followed by a detailed formulation of Kohn-Sham DFT and the derivation of Hellmann-Feynman forces via pseudopotential. Chapter 3 introduces the theoretical constituents of DFPT: the *2n+1 theorem* of Hylleraas and the variational perturbation development of Sinanoğlu with an adaptation to the DFT Hamiltonian [17, 18]. Upon these foundations, it derives the second-order energy perturbation and the corresponding *Sternheimer equation* for the energy definition in Chapter 2. Chapter 3 concludes by discussing the gauge choices necessary for calculating non-degenerate and degenerate orbitals in the Sternheimer equation. Chapter 4 details the weak formulation and discrete equations of FEM. It continues with an overview of the NURBS discretization. Chapter 4 ends by comparing DFT and DFPT schemes such as correlations in problems, inputs and outputs. Chapter 5 demonstrates accuracy of the DFPT formulation and the FEM framework by comparing with contemporary codes and FD schemes. These are categorized for the LDA and GGA exchange-correlation functionals with respect to the orbital degeneracies. The convergence rates of energy, force and IFC

elements are monitored. The last chapter gives a study summary. Finally, it mentions theoretical and computational future work recommendations.

Appendix A derives and validates the DFPT formulation for all-electron framework. At first, it finds the second-order energy perturbation with respect to atomic positions and the corresponding Sternheimer equations. Then, the strong and weak forms are expressed which is followed by the NURBS discretization. The all-electron formulation is also validated with FD. Appendix B outlines several manipulations available for variational principles, especially FEM. Appendix C explains the FD scheme used in validation of DFPT.

Chapter 2

Density Functional Theory

Density Functional Theory (DFT) is a non-interacting electron density-based approach to electronic structure calculations that encapsulates quantum effects into an *exchange-correlation functional*, avoiding the exponential scaling of the many-body Schrödinger equation. Considering *the Born-Oppenheimer approximation*, which neglects the kinetic energy contributions of nuclei due to their larger mass relative to electrons, the Schrödinger Hamiltonian consists of the electron kinetic energy, electron-electron interactions, and electron-nucleus interactions [19]. The Hamiltonian does not include contributions of the nucleus-nucleus interaction due to its electron density independence.

An early model of DFT, *the Thomas-Fermi*, defined the kinetic energy based on a non-interacting uniform electron gas and the electron-electron interactions via the classical Coulomb repulsion. Subsequently, *the Thomas-Fermi-Dirac* model introduced an exchange-energy term to account for the effective same-spin electron repulsion due to the quantum mechanical indistinguishability, considering the *Pauli exclusion principle* [20].

The *first Hohenberg-Kohn theorem* assures a one-to-one correspondence between the external potential $v_{ext}(\mathbf{r})$ and the electron density $\rho(\mathbf{r})$, up to an additive constant. Since the number of electrons and the external potential $v_{ext}(\mathbf{r})$

determines the Hamiltonian, $\rho(\mathbf{r})$ is enough to fix both the ground-state wavefunctions and all observables. On the other hand, the *second Hohenberg-Kohn theorem* applies the *variational principle* to the electron density-based energy definition. It proves that any trial electron density minimizing the energy functional is the ground state electron density. These are what ensure the many-body true wavefunction based Schrödinger equation avoidable.

The Kohn-Sham DFT formulation introduces a framework of non-interacting particles that reproduces the ground-state density of the interacting system. While it models the dominant part of the true kinetic energy exactly, the unknown parts of kinetic energy and other many-body quantum effects in the electron-electron interactions are embedded into the exchange-correlation functional. The Kohn-Sham DFT approach is exact in principle but approximate in practice due to the hidden nature of the exchange-correlation functional.

2.1 Kohn-Sham density functional formulation

Energy formulation in the Kohn-Sham DFT for non-periodic systems can be applied to both all-electron and pseudopential calculations. Pseudopotentials reduce computational complexity by canceling the core electrons with the positively charged nucleus and proposing a potential approximate to the true effects [21]. Therefore, only the valence electron densities are needed in the *self-consistent field (SCF)* iterations and any further calculations.

In the pseudopotential framework, each nucleus is represented by $I \in \{1, \ldots, N\}$ with a charge Z_I and position vector $\mathbf{R}_I$. Any point in the domain is positioned with the vector $\mathbf{r}$. The relative position vector of a nucleus is assigned as $\mathbf{r}_I = \mathbf{r} - \mathbf{R}_I$. The electron density $\rho(\mathbf{r})$ is determined through a set of square-integrable orthonormal states $\psi_\alpha(\mathbf{r})$ and occupancy factors $f_\alpha \in [0, 1]$ with

$$\rho(\mathbf{r}) = 2 \sum_{\alpha}^{occ/2} f_\alpha \langle \psi_\alpha(\mathbf{r}) \psi_\alpha(\mathbf{r}) \rangle. \tag{2.1}$$

The Kohn-Sham formulation requires $\rho(\mathbf{r}) \geq 0$ at any point in the domain and it invokes constraints $\sum_{\alpha}^{occ} f_\alpha = N$ and $\langle \rho(\mathbf{r}) \rangle = N$. $\langle a \rangle$ represents integration over all coordinates, which is equivalent to $\int \int \cdots a \cdots \, d^3\mathbf{r}' \, d^3\mathbf{r}$. Both are used interchangeable in this thesis. Additionally, variables' dependencies on domains can be represented interchangeable as well, such as $\rho = \rho(\mathbf{r})$ and $\rho' = \rho(\mathbf{r}')$ whereas $b = b(\mathbf{r})$ and $b' = b(\mathbf{r}')$.

The ground state Mermin free energy $\tilde{E}$ of a non-periodic system is defined as,

$$\tilde{E} = T_s + E_{NL} + E_{xc} + E_C - \theta S_s, \tag{2.2}$$

with the constant temperature θ and entropy of a non-interacting Kohn-Sham reference system S_s. With the *Laplacian* ∇^2, the kinetic energy contribution T_s is defined as,

$$T_s = -\sum_{\alpha}^{occ/2} f_\alpha \langle \psi_\alpha(\mathbf{r})(\nabla^2)\psi_\alpha(\mathbf{r}) \rangle. \tag{2.3}$$

In this work, the separable norm-conserving formulation of pseudopotentials are preferred [22, 23]. This form contains both local and nonlocal contributions. The local contributions are embedded into a classical electrostatic problem for electron-nucleus and nucleus-nucleus interactions. On the other hand, the nonlocal pseudopotential' s energy contribution E_{NL} is formulated as,

$$\begin{aligned} E_{NL} &= 2 \sum_{\alpha}^{occ/2} f_\alpha \langle \psi_\alpha(\mathbf{r}) \hat{v}_{NL}(\mathbf{r}_I, \mathbf{r}_I') \psi_\alpha(\mathbf{r}') \rangle, \\ &= 2 \sum_{\alpha}^{occ/2} \sum_{I} \sum_{l=0}^{l_I} \sum_{m=-l}^{l} \sum_{k=1}^{k_I} f_\alpha \langle \psi_\alpha(\mathbf{r}) \lambda_{Ik}^{lm}(\mathbf{r}_I) \rangle h_{Ik}^{l} \langle \lambda_{Ik}^{lm}(\mathbf{r}_I') \psi_\alpha(\mathbf{r}') \rangle \end{aligned} \tag{2.4}$$

where the projectors are defined as $\lambda_{Ik}^{lm}(\mathbf{r}_I) = Y_{lm}(\hat{\mathbf{r}}_I) p_{Ik}^{l}(r_I)$ for spherical harmonics Y_{lm} and k_I radial components p_{Ik}^{l} based on Gaussian forms.

The exchange-correlation energy contribution E_{xc} is modeled with both the *local density approximation* (LDA) and *generalized gradient approximation* (GGA) [4, 24]. In this work, only GGA is compatible with the *non-linear core-correction (NLCC)* where ρ_c is a partial core charge density obtained with radial

ionic contributions $\kappa_I(r_I)$ which is also based on Gaussian forms [22]. The core-corrected electron density and density gradient dependent term needed for GGA, are defined as $\bar{\rho} = \rho + \rho_c$ and $\bar{\sigma} = |\nabla\bar{\rho}|^2$, where $\rho_c(\mathbf{r}_I) = \sum_I \kappa_I(\mathbf{r}_I)$,

$$E_{xc} = \langle \bar{\rho} e_{xc}(\bar{\rho}, \bar{\sigma}) \rangle. \tag{2.5}$$

In the LDA, the electron density core-correction becomes arbitrary and the $\bar{\sigma}$ dependence of the exchange-correlation functional falls.

The LIBXC library offers energy definitions in two different ways,

$$E_{xc} = \langle \epsilon_{xc}(\mathbf{r}) \rangle = \langle e_{xc}\rho(\mathbf{r}) \rangle, \tag{2.6}$$

where e_{xc} is energy density per unit particle and ϵ_{xc} is energy density [25]. To find the unperturbed energy, e_{xc} is used whereas the perturbed energies are defined upon ϵ_{xc}.

On the other hand, the exchange-correlation potential, v_{xc}, is defined for the Kohn-Sham Hamiltonian within the variation $\delta E_{xc} = \langle v_{xc}\delta\rho \rangle$. Given the definition in equation (2.6), the exchange-correlation potential for LDA is derived as,

$$v_{xc} = \frac{\partial e_{xc}}{\partial \rho}\rho + e_{xc}, \tag{2.7}$$

while the GGA potential takes the form of,

$$v_{xc} = \frac{\partial e_{xc}}{\partial \bar{\rho}}\bar{\rho} + e_{xc} + 2\bar{\rho}\frac{\partial e_{xc}}{\partial \bar{\sigma}}\nabla\bar{\rho}\cdot\nabla. \tag{2.8}$$

The electrostatic energy contribution E_c covers the Coulomb interactions [26, 27],

$$E_C = E_H + E_{eI} + E_{II}. \tag{2.9}$$

Similar to the exchange-correlation potentials, the electrostatic potential $\hat{v}_C$ is also defined within the variation $\delta E_C = \langle \hat{v}_C \delta\rho \rangle$ so that $\hat{v}_C = \frac{\delta E_C}{\delta\rho}$. The electrostatic energy E_C is determined via potentials calculated with the Poisson equation,

$$\begin{aligned} E_C &= \langle (\rho + b)\hat{v}_C \rangle, \\ -\frac{1}{4\pi}\nabla^2\hat{v}_C &= \rho + b, \end{aligned} \tag{2.10}$$

where, nuclei charges are determined by $b(\mathbf{r}) = \sum_{I=1}^{M} \beta_I(r_I)$ for the number of atoms M. Similarly, the nuclei-nuclei interaction E_{II} is defined using the external potential $\hat{v}_{ext}$ found by separately solving Poisson equations for each $\beta_I(r_I)$ where $\hat{v}_{ext} = \sum_{I=1}^{M} \hat{v}_I$. In order to prevent the self-interaction, a corresponding term is subtracted,

$$E_{II} = \frac{1}{2}\langle b\hat{v}_{ext}\rangle - \frac{1}{2}\sum_{I=1}^{M}\langle \beta_I \hat{v}_I\rangle. \tag{2.11}$$

The hat over the potential terms indicate the discrete calculations. If the same term is used without hat, it represents the exact counterpart.

The entropy term appearing in the Mermin free energy is applied for the non-interacting Kohn-Sham reference system,

$$S_s = -2k\sum_{\alpha}^{occ/2}[f_\alpha ln f_\alpha + (1-f_\alpha)ln(1-f_\alpha)], \tag{2.12}$$

where k is the Boltzmann constant [21].

All the Kohn-Sham DFT calculations are implemented on the FEM numerical settings. For accurate solutions, computationally demanding DFT calculations require high FEM *degree of freedom (DOF)* that results in high run time. In order to achieve more accurate solutions with a low DOF, the numerical electron-nucleus and nucleus-nucleus interactions' external potential related terms can be replaced with analytical counterparts. The discrete electron-nuclei interaction term appearing in the electrostatic energy can be changed by the exact one as [11],

$$E^{\Delta}_{e^-I} = \langle \rho(v_{ext} - \hat{v}_{ext})\rangle. \tag{2.13}$$

Also the nuclei-nuclei interaction term can be revised by,

$$E^{\Delta}_{II} = -(\frac{1}{2}\langle b\hat{v}_{ext}\rangle - \frac{1}{2}\sum_{I=1}^{M}\langle \hat{\beta}_I \hat{v}_I\rangle) + E^{\delta}_{II}. \tag{2.14}$$

The exact nuclei-nuclei interaction energy E^{δ}_{II} becomes a point-charge Z as the ball charges around nuclei do not overlap.

$$E^{\delta}_{II} = \frac{1}{2}\sum_{I=1}^{M}\sum_{J\neq I}^{M}\frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|} \tag{2.15}$$

The energy expression is finalized after the exact-discrete change as,

$$E = \tilde{E} + E^{\Delta}_{e^- I} + E^{\Delta}_{II}. \tag{2.16}$$

Under the normalization constraint $\langle \psi_\alpha \psi_\alpha \rangle = 1$ through an orbital free energy ϵ_α as a corresponding Lagrange multiplier, variation of the energy expression (2.16) with respect to ψ_α delivers the canonical Kohn-Sham equation,

$$(-\frac{1}{2}\nabla^2 + v_c + v_{NL} + v_{xc})\psi_\alpha = \epsilon_\alpha \psi_\alpha. \tag{2.17}$$

The DFT scheme ensures an absolute minimization by iteratively improving the wavefunction. However, this monotonic minimization property strictly holds only when the electrostatic potential is calculated exactly, where the exact electrostatic potential, v_c, consists of both the Hartree potential, $v_H = \int \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\, d^3\mathbf{r}'$, and the external potential, v_{ext}. However, in the FEM framework, direct evaluation of the Hartree potential is problematic due to its long-range nature, $(1/r)$, and divergent behavior near the core. As a result, the electrostatic contribution must be computed via the Poisson equation, as formulated in (2.10). Therefore, the SCF iteration scheme within FEM delivers a saddle-point problem where the total energy is minimized with respect to the Kohn-Sham orbitals whereas it is maximized with respect to the electrostatic potential in the Poisson problem.

2.2 The Hellmann-Feynman forces

The *Hellmann-Feynman* theorem proves that knowledge of the first-order wavefunction perturbations is redundant for force calculation in the case of a self-adjoint Hamiltonian $\hat{H}$ with a normalization condition on the wavefunction $\langle \psi^* \psi \rangle = 1$ [28]. A self-adjoint Hamiltonian has the property of $\hat{H}\psi = \epsilon\psi$ and $\hat{H}\psi^* = \epsilon\psi^*$. For the energy definition,

$$U = \langle \psi^* \hat{H} \psi \rangle, \tag{2.18}$$

the first-order derivative equals to,

$$\frac{\partial U}{\partial R} = \langle \frac{\partial \psi^*}{\partial R} \hat{H} \psi \rangle + \langle \psi^* \frac{\partial \hat{H}}{\partial R} \psi \rangle + \langle \psi^* \hat{H} \frac{\partial \psi}{\partial R} \rangle. \tag{2.19}$$

Because the eigenvalue is independent of the position, it can be trivially moved out of the integral,

$$\frac{\partial U}{\partial R} = \langle \psi^* \frac{\partial \hat{H}}{\partial R} \psi \rangle + \epsilon \langle \frac{\partial \psi^*}{\partial R} \psi + \psi^* \frac{\partial \psi}{\partial R} \rangle. \tag{2.20}$$

The second term in (2.20) corresponds to the first-order derivative of the normalization condition $\langle \psi^* \psi \rangle = 1$, which also represents change in the number of electrons in a closed system. Relying on the constant number of electrons, the second term is canceled out and the force term becomes,

$$F = -\frac{\partial U}{\partial R} = -\langle \psi^* \frac{\partial \hat{H}}{\partial R} \psi \rangle. \tag{2.21}$$

When deriving the force expression for (2.16), explicit dependence on the ionic positions is searched. Therefore, the kinetic energy and Hartree potential do not contribute to the force. In order to benefit from (2.21), a Kohn-Sham Hamiltonian is defined as,

$$\hat{H} = -\frac{1}{2}\nabla^2 + \hat{v}_{NL} + e_{xc} + \frac{1}{2}\hat{v}_H + \hat{v}_{ext}. \tag{2.22}$$

In (2.22), the Hartree potential $\hat{v}_H$ and $\hat{v}_{ext}$ are separated contrary to (2.10) even though they are solved in a single Poisson problem. It is not only unnecessary to calculate $\hat{v}_C$ because the Hartree part does not contribute to the force term but also it is impossible to calculate due to the first-order density perturbation need for the Poisson calculation. Then, the first-order derivative of Hamiltonian $\hat{H}^{(\mathbf{R}_I)}$ can be derived as,

$$\hat{H}^{(\mathbf{R}_I)} = \hat{v}_{NL}^{(\mathbf{R}_I)} + e_{xc}^{(\mathbf{R}_I)} + \hat{v}_{ext}^{(\mathbf{R}_I)}. \tag{2.23}$$

The energy definition (2.16) contains also electron density independent terms, which do not appear in the Hamiltonian derivative (2.23). Therefore, it is necessary to calculate the explicit NLCC contribution $\langle \rho_c(\mathbf{r}_I) e_{xc}(\bar{\rho}, \bar{\sigma}) \rangle$ and nuclei-nuclei interaction terms. Including these extra terms, the revised energy definition

for (2.16) is written as,

$$
\begin{aligned}
E = 2 \sum_{\alpha}^{occ/2} f_\alpha \langle \psi_\alpha^* \hat{H} \psi_\alpha \rangle + \langle \rho_c e_{xc}[\bar{\rho}, \bar{\sigma}] \rangle - \theta S_s \\
+ \langle \rho (v_{ext} - \hat{v}_{ext}) \rangle - (\frac{1}{2} \langle b \hat{v}_{ext} \rangle - \frac{1}{2} \sum_I \langle \hat{\beta}_I \hat{v}_I \rangle) + E_{II}^{\delta}.
\end{aligned}
\tag{2.24}
$$

Considering explicit ionic position $\mathbf{R}_I$ dependence while taking derivatives, the final force expression on the ion I is,

$$
\begin{aligned}
\mathbf{F}_I = -4 \sum_{\alpha}^{occ/2} \sum_{l=0}^{l_I} \sum_{m=-l}^{l} \sum_{k=1}^{k_I} f_\alpha \langle \psi_\alpha^* \lambda_{Ik}^{lm} \rangle h_{Ik}^l \langle \lambda_{Ik}^{lm} \nabla \psi_\alpha \rangle \\
- \langle \bar{\rho} \left(\rho_{c,\mathbf{R}_I} \frac{\partial e_{xc}}{\partial \bar{\rho}} + 2 \nabla \rho_{c,\mathbf{R}_I} \frac{\partial e_{xc}}{\partial \bar{\sigma}} \cdot \nabla \bar{\rho} \right) \rangle - \langle \rho_{c,\mathbf{R}_I} e_{xc}[\bar{\rho}, \bar{\sigma}] \rangle \\
- \langle \rho v_{I,\mathbf{R}_I} \rangle + \sum_{J \neq I} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|^3} (\mathbf{R}_I - \mathbf{R}_J),
\end{aligned}
\tag{2.25}
$$

where $\rho_{c,\mathbf{R}_I}(\mathbf{r}_I)$ is equivalent to $\kappa_{I,\mathbf{R}_I}(\mathbf{r}_I)$ [11]. Several elimination steps during the derivation can be addressed: the discrete external potential term $\langle \rho \hat{v}_{ext} \rangle$ cancels out the electron-nuclei interaction term in the Hamiltonian. Similarly, the discrete nuclei-nuclei and self-interaction terms in (2.24) are dropped due to cancellation with the exact nuclei-nuclei interaction term in the energy definition.

Chapter 3

Density Functional Perturbation Theory

DFPT provides a systematic framework for predicting various material properties by utilizing the linear response of the electron density to perturbations such as atomic displacements, electric fields, and strains in different orders [29]. First-order energy perturbations yield forces, polarizations and stresses, while second-order energy perturbations give force-constant matrix, dielectric susceptibility, elatic constants, Born-effective charge, internal strain, and piezoelectric response [30].It also allows to derive higher-order perturbations.

A $2n + 1^{st}$ order energy perturbation plainly includes wavefunction perturbations up to $2n+1^{st}$ order without any manipulation. This wavefunction perturbation order can be deduced to n^{th} order through imposing *the 2n+1 theorem* [17]. It proves a $2n + 1^{st}$ order energy perturbation with $2n + 1^{st}$ order wavefunction is equivalent to the one calculated with n^{th} order wavefunction.

Unperturbed wavefunctions and corresponding eigenenergies are determined through the Kohn-Sham equation derived via variation of the energy under constraints with respect to the wavefunction, which is quadratically dependent on the wavefunction. On the other hand, each order wavefunction perturbation is

calculated via variation of the even-order energy perturbations with respect to the corresponding wavefunction perturbation [18]. Briefly, an n^{th} order wavefunction perturbation is sufficient to calculate a $2n+1^{st}$ and $2n^{th}$ order energy perturbations whereas an n^{th} order wavefunction perturbation is only available by seeking the stationarity for variation of a $2n^{th}$ order energy perturbation. The governing equations, called *Sternheimer equations*, are solved within SCF iterations [31].

3.1 *2n+1* theorem

Perturbations of any operator, state, or value can be expressed as a power series in a small parameter λ [17],

$$X = X^{(0)} + \lambda X^{(1)} + \lambda^2 X^{(2)} + \lambda^3 X^{(3)} + ..., \tag{3.1}$$

where $X^{(n)} = \frac{1}{n!}\frac{d^n X}{d\lambda^n}$. Given the *Kohn-Sham Hamiltonian* $\hat{H}$, the corresponding *energy eigenvalue* ϵ_α for the *eigenstate* $|\psi_\alpha>$ satisfies the Kohn-Sham (KS) equation,

$$\hat{H}\left|\psi_\alpha\right\rangle = \epsilon_\alpha \left|\psi_\alpha\right\rangle, \tag{3.2}$$

subject to the orthonormality condition,

$$\left\langle\psi_\alpha\psi_\beta\right\rangle = \delta_{\alpha\beta}. \tag{3.3}$$

Accordingly, equations (3.2) and (3.3) are expanded as follows,

$$\begin{aligned}&\left[\hat{H}^{(0)} + \lambda\hat{H}^{(1)} + \lambda^2\hat{H}^{(2)} + \lambda^3\hat{H}^{(3)} + ...\right]\left[\left|\psi_\alpha^{(0)}\right\rangle + \lambda\left|\psi_\alpha^{(1)}\right\rangle + \lambda^2\left|\psi_\alpha^{(2)}\right\rangle + ...\right] \\ &= \left[\epsilon_\alpha^{(0)} + \lambda\epsilon_\alpha^{(1)} + \lambda^2\epsilon_\alpha^{(2)} + \lambda^3\epsilon_\alpha^{(3)} + ...\right]\left[\left|\psi_\alpha^{(0)}\right\rangle + \lambda\left|\psi_\alpha^{(1)}\right\rangle + \lambda^2\left|\psi_\alpha^{(2)}\right\rangle + ...\right],\end{aligned} \tag{3.4}$$

$$\left[\left\langle\psi_\alpha^{(0)}\right| + \lambda\left\langle\psi_\alpha^{(1)}\right| + \lambda^2\left\langle\psi_\alpha^{(2)}\right| + ...\right]\left[\left|\psi_\alpha^{(0)}\right\rangle + \lambda\left|\psi_\alpha^{(1)}\right\rangle + \lambda^2\left|\psi_\alpha^{(2)}\right\rangle + ...\right] = 1. \tag{3.5}$$

At first glance, expanding the Hamiltonian might seem counterintuitive to those familiar with the quantum perturbation theory, where the Hamiltonian is considered as the source of perturbation. The expectation would be to expand energy eigenvalues and eigenstates for the perturbed Hamiltonian. However, in DFT, also the KS Hamiltonian depends on the perturbed eigenstates and eigenvalues,

and thus it must be expanded in λ. By grouping terms of equal order in λ, the equations (3.4) and (3.5) can be formulized as,

$$\sum_{j=0}^{i}(\hat{H}-\epsilon_\alpha)^{(j)}\left|\psi_\alpha^{(i-j)}\right\rangle = 0, \tag{3.6}$$

$$\sum_{j=0}^{i}\langle\psi_\alpha^{(j)}\psi_\alpha^{(i-j)}\rangle = 0, i \neq 0. \tag{3.7}$$

A similar expansion can be derived by premultiplying equation (3.2) with $<\psi_\alpha|$,

$$\langle\psi_\alpha(\hat{H}-\epsilon_\alpha)\psi_\alpha\rangle = 0, \tag{3.8}$$

and it gives the identity of,

$$\sum_{j=0}^{i}\sum_{k=0}^{i}\langle\psi_\alpha^{(j)}(\hat{H}-\epsilon_\alpha)^{(i-j-k)}\psi_\alpha^{(k)}\rangle = 0. \tag{3.9}$$

For second-order variation, the triangular structure similar to the formulation in [32] is obtained as:

$$\begin{aligned}0 = &\langle\psi_\alpha^{(2)}(\hat{H}^{(0)}-\epsilon_\alpha^{(0)})\psi_\alpha^{(0)}\rangle \\ &+\langle\psi_\alpha^{(1)}(\hat{H}^{(1)}-\epsilon_\alpha^{(1)})\psi_\alpha^{(0)}\rangle + \langle\psi_\alpha^{(1)}(\hat{H}^{(0)}-\epsilon_\alpha^{(0)})\psi_\alpha^{(1)}\rangle \\ &+\langle\psi_\alpha^{(0)}(\hat{H}^{(2)}-\epsilon_\alpha^{(2)})\psi_\alpha^{(0)}\rangle + \langle\psi_\alpha^{(0)}(\hat{H}^{(1)}-\epsilon_\alpha^{(1)})\psi_\alpha^{(1)}\rangle + \langle\psi_\alpha^{(0)}(\hat{H}^{(0)}-\epsilon_\alpha^{(0)})\psi_\alpha^{(2)}\rangle.\end{aligned} \tag{3.10}$$

This expression corresponds to the second derivative of the Kohn–Sham equation with respect to atomic positions, divided by two. Using identity (3.9), the first row and the last column of equation (3.10) can be directly eliminated. By applying identity (3.7), the first-order eigenvalue perturbations $\epsilon_\alpha^{(1)}$ can be canceled out, yielding the second-order eigenvalue perturbation $\epsilon_\alpha^{(2)}$ as,

$$\begin{aligned}\epsilon_\alpha^{(2)} = &\langle\psi_\alpha^{(1)}\hat{H}^{(1)}\psi_\alpha^{(0)}\rangle + \langle\psi_\alpha^{(1)}(\hat{H}^{(0)}-\epsilon_\alpha^{(0)})\psi_\alpha^{(1)}\rangle \\ &+\langle\psi_\alpha^{(0)}\hat{H}^{(2)}\psi_\alpha^{(0)}\rangle + \langle\psi_\alpha^{(0)}\hat{H}^{(1)}\psi_\alpha^{(1)}\rangle.\end{aligned} \tag{3.11}$$

As observed in equation (3.11), the second-order variation of the eigenvalue explicitly depends on the zeroth and first-order eigenstate perturbations. This aligns with the essence of the $2n+1$ theorem, which suggests that to compute a perturbation of order $2n$ or $2n+1$, knowledge of wavefunction perturbations up to order n suffices. Therefore, the first-order wavefunction calculation is enough to calculate the second-order energy perturbation.

3.2 Variational perturbation orders

Perturbation equations for the Kohn-Sham DFT are derived using variational methods similar to the Rayleigh-Schrödinger perturbation theory [33]. However, it changes due to a fundamental difference originating from the Hamiltonian definition. The non-relativistic Schrödinger Hamiltonian is independent of the wavefunction. Instead, it consists solely of kinetic and potential energies, typically Laplacians and Coulomb interactions, while quantum effects arise through the wavefunction solutions. In contrast, the Kohn-Sham Hamiltonian explicitly depends on the electron density, which is a functional of the wavefunctions. As a result, the Kohn-Sham equation is required to be solved via SCF iterations instead of one-shot eigenvalue equation. Therefore, this nonlinearity makes variational methods natural for deriving perturbations of the Kohn-Sham DFT.

Solving the Kohn-Sham equation $\hat{H}\psi = \epsilon\psi$ for normalized eigenfunction is equivalent to searching for ψ via,

$$E = \langle\psi\hat{H}\psi\rangle/\langle\psi\psi\rangle. \tag{3.12}$$

For the exact ground state wavefunction and corresponding energy, any trial wavefunction $\tilde{\psi}$ results in,

$$E \leq \langle\tilde{\psi}\hat{H}\tilde{\psi}\rangle/\langle\tilde{\psi}\tilde{\psi}\rangle. \tag{3.13}$$

Once the trial wavefunction $\tilde{\psi} = \sum_{n=0}^{\infty}\phi^{(n)}$ and Hamiltonian $\hat{H} = \sum_{n=0}^{\infty}\hat{H}^{(n)}$ are plugged into equation (3.13), the following relation appears,

$$E \leq E^{(0)} + E^{(1)} + E^{(2)} + E^{(R)} \tag{3.14}$$

where, $E^{(0)}$ is the zeroth-order energy,

$$\begin{aligned} E^{(1)} &= \langle\phi^{(0)}\hat{H}^{(1)}\phi^{(0)}\rangle, \\ E^{(2)} &= 2\langle\phi^{(0)}(\hat{H}^{(1)} - E^{(1)})\phi^{(1)}\rangle + \langle\phi^{(1)}(\hat{H}^{(0)} - E^{(0)})\phi^{(1)}\rangle + \langle\phi^{(0)}(\hat{H}^{(2)} - E^{(2)})\phi^{(0)}\rangle. \end{aligned} \tag{3.15}$$

In this variational formulation, λ terms disappear because it is not an expansion but an exact representation of the contribution at that order. For equation (3.15), the remainder energy $E^{(R)}$ is the error after the second-order energy correction to

the true ground-state energy. To calculate the remainder, higher order derivatives are required. The second-order energy perturbation is explicitly dependent at most on the first-order wavefunction $\phi^{(1)}$. It meets the variational (extrema) property by being quadratic in $\phi^{(1)}$. Therefore, $\phi^{(1)}$ minimizing the second-order energy perturbation will be the solution.

The third-order energy perturbation, $E^{(3)}$, which appears in the remainder of equation (3.14), is derived as,

$$\begin{aligned} E^{(3)} &= \langle\phi^{(0)}\hat{H}^{(3)}\phi^{(0)}\rangle + 2\langle\phi^{(1)}(\hat{H}^{(2)} - E^{(2)})\phi^{(0)}\rangle \\ &\quad + 2\langle\phi^{(2)}(\hat{H}^{(1)} - E^{(1)})\phi^{(0)}\rangle + 2\langle\phi^{(2)}(\hat{H}^{(0)} - E^{(0)})\phi^{(1)}\rangle. \end{aligned} \tag{3.16}$$

The variational property of $E^{(2)}$ with respect to $\phi^{(1)}$ has been proven, and the 2n+1 theorem assures that $E^{(3)}$ can be computed without requiring wavefunction perturbation beyond $\phi^{(1)}$. However, $E^{(3)}$ cannot be used to determine $\phi^{(1)}$, as it is not quadratic in $\phi^{(1)}$. Calculations for further perturbations suggest an observation that even-order energy corrections have variational properties, whereas odd-order corrections do not.

3.3 Sternheimer equation

General form of the Kohn-Sham energy definition that is going to be studied in this thesis contains kinetic energy, nonlocal pseudopotential energy in the *norm-conserving seperable form*, electrostatic energy, exchange-correlation energy for both *Local Density Approximation (LDA)* and *Generalized Gradient Approximation (GGA)* with *Non-Linear Core Correction (NLCC)*, and self-interaction energy contributions, which is represented as,

$$\begin{aligned} \tilde{E} &= \sum_{\alpha}^{occ}\langle\psi_\alpha(\mathbf{r})\hat{T}(\mathbf{r})\psi_\alpha(\mathbf{r})\rangle + \sum_{\alpha}^{occ}\langle\psi_\alpha(\mathbf{r})\hat{V}_{NL}(\mathbf{r},\mathbf{r}')\psi_\alpha(\mathbf{r}')\rangle \\ &\quad + E_C[\rho+b] + E_{xc}[\rho] - \sum_A \frac{1}{2}\left\langle\beta_A\hat{v}_A\right\rangle. \end{aligned} \tag{3.17}$$

The perturbation calculation is directly available via equation (3.17). However, in order to create a 2n+1 theorem compatible energy formulation, the equation (3.17) is modified to,

$$\begin{aligned}\tilde{E} &= \sum_{\alpha}^{occ} \epsilon_\alpha + E_C[\rho + b] + E_{xc}[\rho] \\ &- \int \frac{\delta E_C[\rho + b]}{\delta(\rho(\mathbf{r}) + b(\mathbf{r}))} \rho(\mathbf{r})\, d^3\mathbf{r} - \int \frac{\delta E_{xc}[\rho]}{\delta \rho(\mathbf{r})} \rho(\mathbf{r})\, d^3\mathbf{r} - \sum_A \frac{1}{2} \left\langle \beta_A \hat{v}_A \right\rangle ,\end{aligned} \tag{3.18}$$

where,

$$\epsilon_\alpha = \langle \psi_\alpha \left[\hat{T} + \hat{V}_{NL} + \frac{\delta E_C[\rho + b]}{\delta(\rho(\mathbf{r}) + b(\mathbf{r}))} + \frac{\delta E_{xc}[\rho]}{\delta \rho(\mathbf{r})} \right] \psi_\alpha \rangle. \tag{3.19}$$

The energy form (3.18) simplifies the perturbation derivation including the orthonormalization condition, given the energy with Lagrange multipliers as [34],

$$\tilde{F} = \tilde{E} - \sum_{\alpha}^{occ} \sum_{\beta}^{occ} \epsilon_{\alpha\beta} (\langle \psi_\alpha \psi_\beta \rangle - \delta_{\alpha\beta}). \tag{3.20}$$

Equation (3.20) is the parallel transport gauge compatible energy form with Lagrange multipliers. This form is applicable to both non-degenerate and degenerate orbitals in any molecular systems. However, only the diagonal gauge formulation is given due to its simplicity in which the multipliers are diagonal. The energy definition upon which the perturbation formulation is developed is determined as,

$$\tilde{F} = \tilde{E} - \sum_{\alpha}^{occ} \epsilon_\alpha (\langle \psi_\alpha \psi_\alpha \rangle - I). \tag{3.21}$$

As a notification, the parallel transport gauge compatible formulation is going to contain off-diagonal occupation factors so that it is also required to be transformed. Then, the DFPT formulation follows the same derivation procedure.

For equation (3.21), the second-order energy perturbation in general form becomes,

$$
\begin{aligned}
\tilde{F}^{(2)} = & \sum_{\alpha} f_{\alpha}[\langle\psi_{\alpha}^{(1)}(\hat{T}^{(1)}+\hat{V}_{NL}^{(1)})\psi_{\alpha}^{(0)}\rangle + \langle\psi_{\alpha}^{(1)}(\hat{H}^{(0)}-\epsilon_{\alpha}^{(0)})\psi_{\alpha}^{(1)}\rangle \\
& + \langle\psi_{\alpha}^{(0)}(\hat{T}^{(2)}+\hat{V}_{NL}^{(2)})\psi_{\alpha}^{(0)}\rangle + \langle\psi_{\alpha}^{(0)}(\hat{T}^{(1)}+\hat{V}_{NL}^{(1)})\psi_{\alpha}^{(1)}\rangle] \\
& + \sum_{\alpha} f_{\alpha}[\langle\psi_{\alpha}^{(1)}\left[\frac{\delta E_C[\rho+b]}{\delta(\rho(\mathbf{r})+b(\mathbf{r}))}\right]^{(1)}\psi_{\alpha}^{(0)}\rangle + \langle\psi_{\alpha}^{(0)}\left[\frac{\delta E_C[\rho+b]}{\delta(\rho(\mathbf{r})+b(\mathbf{r}))}\right]^{(1)}\psi_{\alpha}^{(1)}\rangle] \\
& + \sum_{\alpha} f_{\alpha}\langle\psi_{\alpha}^{(0)}\left[\frac{\delta E_C[\rho+b]}{\delta(\rho(\mathbf{r})+b(\mathbf{r}))}\right]^{(2)}\psi_{\alpha}^{(0)}\rangle \\
& - \left[\int \frac{\delta E_C[\rho+b]}{\delta(\rho(\mathbf{r})+b(\mathbf{r}))}\rho(\mathbf{r})\,d^3\mathbf{r}\right]^{(2)} + [E_C[\rho+b]]^{(2)} \\
& + \sum_{\alpha} f_{\alpha}[\langle\psi_{\alpha}^{(1)}\left[\frac{\delta E_{xc}[\rho]}{\delta\rho(\mathbf{r})}\right]^{(1)}\psi_{\alpha}^{(0)}\rangle + \langle\psi_{\alpha}^{(0)}\left[\frac{\delta E_{xc}[\rho]}{\delta\rho(\mathbf{r})}\right]^{(1)}\psi_{\alpha}^{(1)}\rangle] \\
& + \sum_{\alpha} f_{\alpha}\langle\psi_{\alpha}^{(0)}\left[\frac{\delta E_{xc}[\rho]}{\delta\rho(\mathbf{r})}\right]^{(2)}\psi_{\alpha}^{(0)}\rangle - \left[\int \frac{\delta E_{xc}[\rho]}{\delta\rho(\mathbf{r})}\rho(\mathbf{r})\,d^3\mathbf{r}\right]^{(2)} + [E_{xc}[\rho]]^{(2)} \\
& - \left[\sum_{A}\frac{1}{2}\langle\beta_A\hat{v}_A\rangle\right]^{(2)}.
\end{aligned}
\tag{3.22}
$$

In equation (3.22) and the following equations, the perturbations on the non-local contribution functions λ are transferred to the other elements using the integration by parts and gradient theorems (see Appendix B). It benefits from the equivalence of derivatives so that it could change the atomic position perturbations to gradients when they operate on $\mathbf{r}_I$.

Relying on Chapter 3.2, the second-order energy functional is stationary with respect to the first-order wavefunctions, thereby enabling its treatment via variational optimization schemes. Therefore, the Sternheimer equation is derived by taking derivative with respect to the first-order wavefunction. For clarification, the Sternheimer equation is held in the strong form of FEM, attributing $\delta\psi_{\alpha}^{(1)}$ as a test function, which is mentioned in the subsequent sections. Considering that equation (3.22) is the duplicate of second-order derivative with division by 2, it takes the form of,

$$
\begin{aligned}
\frac{\delta}{\delta\psi_\alpha^{(1)}}\left[\frac{d^2\tilde{F}}{d\lambda d\lambda}\right] &= 4f_\alpha\langle\delta\psi_\alpha^{(1)}(\hat{H}^{(0)}-\epsilon_\alpha^{(0)})\psi_\alpha^{(1)}\rangle \\
&+ 2f_\alpha\langle\delta\psi_\alpha^{(1)}\hat{T}^{(1)}\psi_\alpha^{(0)}\rangle + 2f_\alpha\langle\psi_\alpha^{(0)}\hat{T}^{(1)}\delta\psi_\alpha^{(1)}\rangle \\
&+ 2f_\alpha\int\nabla\delta\psi_\alpha^{(1)}\lambda_J^{(0)}(\mathbf{r})\,d^3\mathbf{r}\int\lambda_J^{*(0)}(\mathbf{r}')\psi_\alpha^{(0)}(\mathbf{r}')\,d^3\mathbf{r}' \\
&+ 2f_\alpha\int\delta\psi_\alpha^{(1)}\lambda_J^{(0)}(\mathbf{r})\,d^3\mathbf{r}\int\lambda_J^{*(0)}(\mathbf{r}')\nabla\psi_\alpha^{(0)}(\mathbf{r}')\,d^3\mathbf{r}' \\
&+ 2f_\alpha\int\nabla\psi_\alpha^{*(0)}(\mathbf{r})\lambda_I^{(0)}(\mathbf{r})\,d^3\mathbf{r}\int\lambda_I^{*(0)}(\mathbf{r}')\delta\psi_\alpha^{(1)}\,d^3\mathbf{r}' \\
&+ 2f_\alpha\int\psi_\alpha^{*(0)}(\mathbf{r})\lambda_I^{(0)}(\mathbf{r})\,d^3\mathbf{r}\int\lambda_I^{*(0)}(\mathbf{r}')\nabla\delta\psi_\alpha^{(1)}\,d^3\mathbf{r}' \\
&+ 4f_\alpha\int\delta\psi_\alpha^{(1)}\left[\int\frac{\delta^2E_{xc}[\rho^{(0)}]}{\delta\rho(\mathbf{r})^2}\rho'^{(1)}\,d^3\mathbf{r}'\right]\psi_\alpha^{(0)}(\mathbf{r})\,d^3\mathbf{r} \\
&+ 4f_\alpha\int\delta\psi_\alpha^{(1)}\left[S_C\right]\psi_\alpha^{(0)}(\mathbf{r})\,d^3\mathbf{r},
\end{aligned}
\tag{3.23}
$$

where the first-order perturbation of electrostatic potential S_C is defined as,

$$
S_C = \int\frac{\delta^2E_C[\rho+b]}{\delta(\rho(\mathbf{r})+b(\mathbf{r}))^2}(\rho'^{(1)}+b'^{(1)})\,d^3\mathbf{r}'. \tag{3.24}
$$

Equation (3.23) is equivalent to negative of the search direction in minimization algorithms [35]. So considering the problem as a nonlinear minimization problem, a solution to any system can be found by using Quasi-Newton methods. On the other hand, the problem can be transformed into a linear system of equations as shown in equation (3.25). The solution can be found using direct or iterative

solvers within SCF iterations.

$$
\begin{aligned}
4f_\alpha\langle\delta\psi_\alpha^{(1)}(\hat{H}^{(0)}-\epsilon_\alpha^{(0)})\psi_\alpha^{(1)}\rangle = &-2f_\alpha\langle\delta\psi_\alpha^{(1)}\hat{T}^{(1)}\psi_\alpha^{(0)}\rangle - 2f_\alpha\langle\psi_\alpha^{(0)}\hat{T}^{(1)}\delta\psi_\alpha^{(1)}\rangle \\
&- 2f_\alpha\int\nabla\delta\psi_\alpha^{(1)}\lambda_J^{(0)}(\mathbf{r})\,d^3\mathbf{r}\int\lambda_J^{*(0)}(\mathbf{r}')\psi_\alpha^{(0)}(\mathbf{r}')\,d^3\mathbf{r}' \\
&- 2f_\alpha\int\delta\psi_\alpha^{(1)}\lambda_J^{(0)}(\mathbf{r})\,d^3\mathbf{r}\int\lambda_J^{*(0)}(\mathbf{r}')\nabla\psi_\alpha^{(0)}(\mathbf{r}')\,d^3\mathbf{r}' \\
&- 2f_\alpha\int\nabla\psi_\alpha^{*(0)}(\mathbf{r})\lambda_I^{(0)}(\mathbf{r})\,d^3\mathbf{r}\int\lambda_I^{*(0)}(\mathbf{r}')\delta\psi_\alpha^{(1)}\,d^3\mathbf{r}' \\
&- 2f_\alpha\int\psi_\alpha^{*(0)}(\mathbf{r})\lambda_I^{(0)}(\mathbf{r})\,d^3\mathbf{r}\int\lambda_I^{*(0)}(\mathbf{r}')\nabla\delta\psi_\alpha^{(1)}\,d^3\mathbf{r}' \\
&- 4f_\alpha\int\delta\psi_\alpha^{(1)}\left[\int\frac{\delta^2E_{xc}[\rho^{(0)}]}{\delta\rho(\mathbf{r})^2}\rho'^{(1)}\,d^3\mathbf{r}'\right]\psi_\alpha^{(0)}(\mathbf{r})\,d^3\mathbf{r} \\
&- 4f_\alpha\int\delta\psi_\alpha^{(1)}\left[S_C\right]\psi_\alpha^{(0)}(\mathbf{r})\,d^3\mathbf{r}
\end{aligned}
\tag{3.25}
$$

In this study, the IFC tensor $\mathbf{C}$ is calculated for proof of concept. It corresponds to the second-order atomic position perturbation of energy. The IFC tensor differs from equation (3.22) by the dropped off kinetic energy perturbation terms because it is independent of atomic positions. Then, the second-order atomic positions perturbation of energy is going to take the form of,

$$
\begin{aligned}
\frac{d^2\tilde{F}}{d\mathbf{R}_I d\mathbf{R}_J} &= 4\sum_{\alpha=1}^{occ/2} f_\alpha \langle \psi_\alpha^{(\mathbf{R}_I)} (\hat{H}^{(0)} - \epsilon_\alpha^{(0)}) \psi_\alpha^{(\mathbf{R}_J)} \rangle \\
&+ 4\sum_{\alpha=1}^{occ/2} f_\alpha \int \nabla \psi_\alpha^{*(\mathbf{R}_I)}(\mathbf{r}) \lambda_J^{(0)}(\mathbf{r})\, d^3\mathbf{r} \int \lambda_J^{*(0)}(\mathbf{r}') \psi_\alpha^{(0)}(\mathbf{r}')\, d^3\mathbf{r}' \\
&+ 4\sum_{\alpha=1}^{occ/2} f_\alpha \int \psi_\alpha^{*(\mathbf{R}_I)}(\mathbf{r}) \lambda_J^{(0)}(\mathbf{r})\, d^3\mathbf{r} \int \lambda_J^{*(0)}(\mathbf{r}') \nabla \psi_\alpha^{(0)}(\mathbf{r}')\, d^3\mathbf{r}' \\
&+ 4\sum_{\alpha=1}^{occ/2} f_\alpha \int \nabla \psi_\alpha^{*(0)}(\mathbf{r}) \lambda_I^{(0)}(\mathbf{r})\, d^3\mathbf{r} \int \lambda_I^{*(0)}(\mathbf{r}') \psi_\alpha^{(\mathbf{R}_J)}(\mathbf{r}')\, d^3\mathbf{r}' \\
&+ 4\sum_{\alpha=1}^{occ/2} f_\alpha \int \psi_\alpha^{*(0)}(\mathbf{r}) \lambda_I^{(0)}(\mathbf{r})\, d^3\mathbf{r} \int \lambda_I^{*(0)}(\mathbf{r}') \nabla \psi_\alpha^{(\mathbf{R}_J)}(\mathbf{r}')\, d^3\mathbf{r}' \\
&+ 4\sum_{\alpha=1}^{occ/2} f_\alpha \int \nabla \psi_\alpha^{*(0)}(\mathbf{r}) \lambda_I^{(0)}(\mathbf{r})\, d^3\mathbf{r} \int \lambda_I^{*(0)}(\mathbf{r}') \nabla \psi_\alpha^{(0)}(\mathbf{r}')\, d^3\mathbf{r}' \delta_{IJ} \\
&- 4\sum_{\alpha=1}^{occ/2} f_\alpha \int \nabla \psi_\alpha^{*(0)}(\mathbf{r}) \nabla \lambda_I^{(0)}(\mathbf{r})\, d^3\mathbf{r} \int \lambda_I^{*(0)}(\mathbf{r}') \psi_\alpha^{(0)}(\mathbf{r}')\, d^3\mathbf{r}' \delta_{IJ} \\
&+ \int \left[\int \frac{\delta^2 E_{xc}[\rho^{(0)}]}{\delta\rho(\mathbf{r})^2} \rho'^{(\mathbf{R}_I)} \delta(\mathbf{r}' - \mathbf{r})\, d^3\mathbf{r}' \right] \rho^{(\mathbf{R}_J)}\, d^3\mathbf{r} \\
&+ \frac{1}{2} \int [S_C] \left(\rho^{(\mathbf{R}_J)} + b^{(\mathbf{R}_J)} \right) d^3\mathbf{r} \\
&+ \frac{1}{2} \int [S_C] \left(\rho^{(\mathbf{R}_I)} + b^{(\mathbf{R}_I)} \right) d^3\mathbf{r} \\
&+ \int \frac{\delta E_C[\rho + b]}{\delta(\rho(\mathbf{r}) + b(\mathbf{r}))} b^{(\mathbf{R}_I, \mathbf{R}_I)} \delta_{IJ}\, d^3\mathbf{r} \\
&- \int \int \frac{\delta \hat{v}_I[\beta_I]}{\delta[\beta_I']} \beta_I'^{(\mathbf{R}_I)}\, d^3\mathbf{r}' \beta_I^{(\mathbf{R}_I)} \delta_{IJ}\, d^3\mathbf{r} - \int \hat{v}_I[\beta_I] \beta_I^{(\mathbf{R}_I, \mathbf{R}_I)} \delta_{IJ}\, d^3\mathbf{r}.
\end{aligned}
\tag{3.26}
$$

The atomic position perturbations on any operator, vector, and value are represented by $X^{(\mathbf{R}_I)}$. The $\delta(\mathbf{r}' - \mathbf{r})$ indicates locality of functional whereas δ_{IJ} means that terms will not contribute to the mixed-nuclei perturbations in equation (3.26).

For equation (3.26), the Sternheimer equation is derived as,

$$\begin{aligned}
8f_\alpha\langle\delta\psi_\alpha^{(\mathbf{R}_I)}(\hat{H}^{(0)}-\epsilon_\alpha^{(0)})\psi_\alpha^{(\mathbf{R}_I)}\rangle &= -4f_\alpha\int\nabla\delta\psi_\alpha^{(\mathbf{R}_I)}\lambda_I^{(0)}\,d^3\mathbf{r}\int\lambda_I^{*(0)}\psi_\alpha^{(0)}\,d^3\mathbf{r}' \\
&\quad - 4f_\alpha\int\delta\psi_\alpha^{(\mathbf{R}_I)}\lambda_I^{(0)}\,d^3\mathbf{r}\int\lambda_I^{*(0)}\nabla\psi_\alpha^{(0)}\,d^3\mathbf{r}' \\
&\quad - 4f_\alpha\int\nabla\psi_\alpha^{*(0)}\lambda_I^{(0)}\,d^3\mathbf{r}\int\lambda_I^{*(0)}\delta\psi_\alpha^{(\mathbf{R}_I)}\,d^3\mathbf{r}' \\
&\quad - 4f_\alpha\int\psi_\alpha^{*(0)}\lambda_I^{(0)}\,d^3\mathbf{r}\int\lambda_I^{*(0)}\nabla\delta\psi_\alpha^{(\mathbf{R}_I)}\,d^3\mathbf{r}' \\
&\quad - 8f_\alpha\int\delta\psi_\alpha^{(\mathbf{R}_I)}\left[\int\frac{\delta^2E_{xc}[\rho^{(0)}]}{\delta\rho(\mathbf{r})^2}\rho'^{(\mathbf{R}_I)}\,d^3\mathbf{r}'\right]\psi_\alpha^{(0)}\,d^3\mathbf{r} \\
&\quad - 8f_\alpha\int\delta\psi_\alpha^{(\mathbf{R}_I)}\,[S_C]\,\psi_\alpha^{(0)}\,d^3\mathbf{r}.
\end{aligned} \tag{3.27}$$

where, the first-order electron density is defined as,

$$\rho(\mathbf{r})^{(\mathbf{R}_I)} = 2\sum_{\alpha}^{occ/2} f_\alpha(\psi_\alpha^*(\mathbf{r})^{(\mathbf{R}_I)}\psi_\alpha(\mathbf{r})^{(0)} + \psi_\alpha^*(\mathbf{r})^{(0)}\psi_\alpha(\mathbf{r})^{(\mathbf{R}_I)}). \tag{3.28}$$

First-order variation of the electrostatic potential (2.10) can be preferred to calculate by observing the second-order energy perturbation with respect to the atomic positions (3.26) and the corresponding Sternheimer equation (3.27).

$$\begin{aligned}
-\frac{1}{4\pi}\nabla^2\hat{v}_C^{(\mathbf{R}_I)} &= \rho^{(\mathbf{R}_I)} + b^{(\mathbf{R}_I)} \\
-\frac{1}{4\pi}\nabla^2 S_C &= \rho^{(\mathbf{R}_I)} + b^{(\mathbf{R}_I)}.
\end{aligned} \tag{3.29}$$

The first-order perturbation formulations in equations (3.26), (3.27), and (3.29) display local pseudocharges up to the second-order perturbation. The local pseudocharges are calculated via Laplacian of the Goedecker' s potential definition [23] in the spherical coordinates under the assumption of spherical symmetry with,

$$b_{loc}(r_A) = -\frac{1}{4\pi}\frac{1}{r_A^2}\frac{d}{dr}\left[r_A^2\frac{d}{dr}V_{loc}(r_A)\right]. \tag{3.30}$$

Given A and B,

$$\begin{aligned}
A &= -\frac{2}{r_A^2}\frac{d}{dr}V_{loc}(r_A) + \frac{2}{r_A}\frac{d^2}{dr^2}V_{loc}(r_A) + \frac{d^3}{dr^3}V_{loc}(r_A), \\
B &= \frac{4}{r_A^3}\frac{d}{dr}V_{loc}(r_A) - \frac{4}{r_A^2}\frac{d^2}{dr^2}V_{loc}(r_A) + \frac{2}{r_A}\frac{d^3}{dr^3}V_{loc}(r_A) + \frac{d^4}{dr^4}V_{loc}(r_A),
\end{aligned} \tag{3.31}$$

the first and second-order perturbations can be derived trivially as,

$$b_{loc}^{(\mathbf{R}_I)}(r_A) = -\frac{1}{4\pi} A \left(-\frac{\mathbf{r}_I}{r_A} \right), \tag{3.32}$$

$$b_{loc}^{(\mathbf{R}_I,\mathbf{R}_J)}(r_A) = -\frac{1}{4\pi} A \left(\frac{1}{r_A}\mathbf{I} - \frac{1}{r_A^3}\mathbf{r}_I \otimes \mathbf{r}_J \right) - \frac{1}{4\pi} B \frac{1}{r_A^2}\mathbf{r}_I \otimes \mathbf{r}_J, \tag{3.33}$$

where $\mathbf{r}_A = \mathbf{r} - \mathbf{R}_A$ and $r_A = |\mathbf{r}_A|$. $\mathbf{r}$ represents any position in the domain while $\mathbf{R}_A$ is position of the atom A. I and J indices in equations (3.32) and (3.33) span the perturbations.

Equations (3.26) and (3.27) use the LDA exchange-correlation functional. In order to switch to the GGA contributions, the exchange-correlation contributions in the zeroth-order Hamiltonian and eigenvalue should be changed. Also the exchange-correlation perturbation term should be switched as follows,

$$\begin{aligned}
&\int \left[\int \frac{\delta^2 E_{xc}[\rho^{(0)}]}{\delta\rho(\mathbf{r})^2} \rho'^{(\mathbf{R}_I)} \delta(\mathbf{r}' - \mathbf{r})\, d^3\mathbf{r}' \right] \rho^{(\mathbf{R}_J)}\, d^3\mathbf{r} \\
&\equiv \int \frac{\partial^2 \epsilon_{xc}}{\partial\bar{\rho}^2} \bar{\rho}^{(\mathbf{R}_I)} \bar{\rho}^{(\mathbf{R}_J)}\, d^3\mathbf{r} + \int \frac{\partial^2 \epsilon_{xc}}{\partial\bar{\rho}\partial\bar{\sigma}} \bar{\rho}^{(\mathbf{R}_I)} \bar{\sigma}^{(\mathbf{R}_J)}\, d^3\mathbf{r} \\
&+ \int \frac{\partial^2 \epsilon_{xc}}{\partial\bar{\sigma}\partial\bar{\rho}} \bar{\sigma}^{(\mathbf{R}_I)} \bar{\rho}^{(\mathbf{R}_J)}\, d^3\mathbf{r} + \int \frac{\partial^2 \epsilon_{xc}}{\partial\bar{\sigma}^2} \bar{\sigma}^{(\mathbf{R}_I)} \bar{\sigma}^{(\mathbf{R}_J)}\, d^3\mathbf{r} \\
&+ 2 \int \frac{\partial \epsilon_{xc}}{\partial\bar{\sigma}} \left[\nabla\bar{\rho}^{(\mathbf{R}_I)} \cdot \nabla\bar{\rho}^{(\mathbf{R}_J)} + \nabla\rho_c^{(\mathbf{R}_I,\mathbf{R}_J)} \cdot \nabla\bar{\rho}^{(0)} \right] d^3\mathbf{r} \\
&+ \int \frac{\partial \epsilon_{xc}}{\partial\bar{\rho}} \rho_c^{(\mathbf{R}_I,\mathbf{R}_J)}\, d^3\mathbf{r}.
\end{aligned} \tag{3.34}$$

In order to convert the Sternheimer equation (3.27) in LDA to GGA, the variation of equation (3.34) with respect to the first wavefunction can be taken and replaced.

The GGA format requires a core correction term to the electron density only in the exchange-correlation related terms. In this work, the Willand's non-linear core correction density, ρ_c, is employed by [22],

$$\rho_c(\mathbf{r}_A) = c_{core} \frac{Z - Z_{nucleus}}{\left(\sqrt{2\pi} r_{core}\right)^3} e^{-\frac{1}{2}\left(\frac{r_A}{r_{core}}\right)^2}. \tag{3.35}$$

Then, the atomic perturbations are derived accordingly as,

$$\rho_c^{(\mathbf{R}_I)}(\mathbf{r}_A) = \rho_c^{(0)}(\mathbf{r}_A) \frac{\mathbf{r}_I}{r_{core}^2}, \tag{3.36}$$

and

$$\rho_c^{(\mathbf{R}_I,\mathbf{R}_J)}(\mathbf{r}_A) = \rho_c^{(0)}(\mathbf{r}_A)\left[\frac{1}{r_{core}^4}\mathbf{r}_I \otimes \mathbf{r}_J - \frac{1}{r_{core}^2}\mathbf{I}\right]. \tag{3.37}$$

Also the density gradient term, $\bar{\sigma}^{(\mathbf{R}_I)}$, and the gradient of the densities can be derived considering the definitions, $\bar{\rho} = \rho + \rho_c$ and $\bar{\sigma} = |\nabla\bar{\rho}|^2$.

The energy perturbation and Sternheimer equation accuracies can be improved with an exact-discrete change similar to the DFT equations (2.13) and (2.14). It relies on the replacement of electron-nucleus and nucleus-nucleus interaction potentials created in the Poisson problem with the exact pseudopotential counterparts. If the same perturbation process is followed, correction to the second-order electron-nucleus contribution becomes,

$$\begin{aligned}
-\langle\rho\hat{v}_{ext}\rangle^{(\mathbf{R}_I,\mathbf{R}_J)} + \langle\rho v_{ext}\rangle^{(\mathbf{R}_I,\mathbf{R}_J)} &= 4\sum_{\alpha=1}^{occ/2} f_\alpha\langle\psi_\alpha^{(\mathbf{R}_I)}\left(v_{ext}^{(0)} - \hat{v}_{ext}^{(0)}\right)\psi_\alpha^{(\mathbf{R}_J)}\rangle \\
&+ \int\left(v_{ext}^{(\mathbf{R}_J)} - \hat{v}_{ext}^{(\mathbf{R}_J)}\right)\rho^{(\mathbf{R}_I)}\,d^3\mathbf{r} \\
&+ \int\left(v_{ext}^{(\mathbf{R}_I)} - \hat{v}_{ext}^{(\mathbf{R}_I)}\right)\rho^{(\mathbf{R}_J)}\,d^3\mathbf{r} \\
&+ \int\left(v_{ext}^{(\mathbf{R}_I,\mathbf{R}_J)} - \hat{v}_{ext}^{(\mathbf{R}_I,\mathbf{R}_J)}\right)\rho^{(0)}\,d^3\mathbf{r}.
\end{aligned} \tag{3.38}$$

While finding the discrete electrostatic energy contribution, each nucleus' interaction with itself is manually subtracted. Also the avoidable second-order perturbations are decreased by the integration by parts and gradient theorem so that the discrete nucleus-nucleus interaction becomes,

$$\begin{aligned}
-\frac{1}{2}\langle b\hat{v}_{ext}\rangle^{(\mathbf{R}_I,\mathbf{R}_J)} + E_{self}^{(\mathbf{R}_I,\mathbf{R}_J)} &= \frac{1}{2}\int\hat{v}_A^{(\mathbf{R}_I)}b^{(\mathbf{R}_I)}\,d^3\mathbf{r}\delta_{IJ} - \frac{1}{2}\int\hat{v}_A^{(\mathbf{R}_I)}\beta_A^{(\mathbf{R}_J)}\,d^3\mathbf{r} \\
&- \frac{1}{2}\int\hat{v}_A^{(\mathbf{R}_J)}\beta_A^{(\mathbf{R}_I)}\,d^3\mathbf{r} + \frac{1}{2}\int\hat{v}_{ext}^{(\mathbf{R}_I)}\beta_A^{(\mathbf{R}_I)}\,d^3\mathbf{r}\delta_{IJ} \\
&+ \int\int\frac{\delta\hat{v}_I[\beta_I]}{\delta[\beta_I']}\beta_I'^{(\mathbf{R}_I)}\,d^3\mathbf{r}'\beta_I^{(\mathbf{R}_I)}\delta_{IJ}\,d^3\mathbf{r} \\
&+ \int\nabla\hat{v}_I[\beta_I]\beta_I^{(\mathbf{R}_I)}\delta_{IJ}\,d^3\mathbf{r}.
\end{aligned} \tag{3.39}$$

The exact nucleus-nucleus interaction perturbations can be decomposed into diagonal and off-diagonal nucleus-nucleus terms. Here, diagonal refers to interactions where the perturbation is with respect to the same nucleus index, $I - I$, while off-diagonal refers to mixed derivatives involving different nucleus indices, $I - J$. It means neither the matrix diagonality nor the nuclei' spatial vector components. The diagonal nucleus-nucleus perturbation is given by,

$$E_{II}^{\delta(\mathbf{R}_I,\mathbf{R}_I)} = \sum_{J=1,J\neq I}^{N} Z_I Z_J \left[3\frac{1}{|\mathbf{R}_I - \mathbf{R}_J|^5} (\mathbf{R}_I - \mathbf{R}_J) \otimes (\mathbf{R}_I - \mathbf{R}_J) - \frac{1}{|\mathbf{R}_I - \mathbf{R}_J|^3}\mathbf{I} \right], \tag{3.40}$$

while the off-diagonal nucleus-nucleus perturbation is formulated as,

$$E_{II}^{\delta(\mathbf{R}_I,\mathbf{R}_J)} = -Z_I Z_J \left[3\frac{1}{|\mathbf{R}_I - \mathbf{R}_J|^5} (\mathbf{R}_I - \mathbf{R}_J) \otimes (\mathbf{R}_I - \mathbf{R}_J) - \frac{1}{|\mathbf{R}_I - \mathbf{R}_J|^3}\mathbf{I} \right]. \tag{3.41}$$

The first-order electron density perturbation dependence is only observed in the electron-nucleus discrete-exact change (3.38). Its contribution to the Sternheimer equation (3.27) is derived as,

$$\begin{aligned} 8f_\alpha \langle \delta\psi_\alpha^{(\mathbf{R}_I)} \left(v_{ext}^{(0)} - \hat{v}_{ext}^{(0)} \right) \psi_\alpha^{(\mathbf{R}_I)} \rangle = & -4f_\alpha \langle \delta\psi_\alpha^{(\mathbf{R}_I)} \left(v_{ext}^{(\mathbf{R}_I)} - \hat{v}_{ext}^{(\mathbf{R}_I)} \right) \psi_\alpha^{(0)} \rangle \\ & - 4f_\alpha \langle \psi_\alpha^{(0)} \left(v_{ext}^{(\mathbf{R}_I)} - \hat{v}_{ext}^{(\mathbf{R}_I)} \right) \delta\psi_\alpha^{(\mathbf{R}_I)} \rangle. \end{aligned} \tag{3.42}$$

Beyond these corrections, the unperturbed eigenenergies $\epsilon_\alpha^{(0)}$, which contribute to both the second-order energy perturbation and the Sternheimer equation, are affected by exact-discrete changes due to the corrections in the Kohn-Sham DFT. It is significant to notice that these exact-discrete changes are preferred only within pseudopotential framework, not for the all-electron.

3.4 Gauge selection

The Sternheimer equation (3.27) is compactly expressed in strong form as,

$$(\hat{H}^{(0)} - \epsilon_\alpha^{(0)})\psi_\alpha^{(1)} = -(\hat{H}^{(1)} - \epsilon_\alpha^{(1)})\psi_\alpha^{(0)}, \tag{3.43}$$

where $\epsilon_\alpha^{(1)} = \langle \psi_\alpha^{(0)} \hat{H}^{(1)} \psi_\alpha^{(0)} \rangle$. The operator on the LHS , $(\hat{H}^{(0)} - \epsilon_\alpha^{(0)})$, is a singular operator when applied on the occupied subspace. Consequently, any linear combination of $\psi_\alpha^{(0)}$ added to $\psi_\alpha^{(1)}$ will also satisfy equation (3.43). This degeneracy is called the *gauge freedom*, and it leads to an infinite number of valid solutions [35].

In order to eliminate redundancy, a gauge condition on the wavefunctions must be imposed [18]. An orthonormality condition for the perturbed wavefunctions is,

$$\langle \psi_\alpha^{*(0)} \psi_\beta^{(1)} \rangle + \langle \psi_\alpha^{*(1)} \psi_\beta^{(0)} \rangle = 0. \tag{3.44}$$

Instead, the stricter condition of $\langle \psi_\alpha^{*(0)} \psi_\beta^{(1)} \rangle = 0$ is forced due to the computational concerns of equation (3.44). On the other hand, projectors are imposed on both sides in order to make the Sternheimer equation occupied-state free. The *parallel transport gauge* associated idempotent projector is defined by,

$$\hat{P}_p = \hat{I} - \sum_{\alpha}^{occ} |\psi_\alpha^{(0)}\rangle\langle\psi_\alpha^{(0)}|. \tag{3.45}$$

As a result, the gauge-fixed Sternheimer equation is derived as,

$$\hat{P}_p(\hat{H}^{(0)} - \epsilon_\alpha^{(0)})\hat{P}_p\psi_\alpha^{(1)} = -\hat{P}_p(\hat{H}^{(1)} - \epsilon_\alpha^{(1)})\psi_\alpha^{(0)}. \tag{3.46}$$

It ensures that both sides of equation (3.46) lie in the unoccupied subspace.

Chapter 4

Numerical Framework

Both the FEM framework and the DFPT scheme involve multiple layers of numerical consideration. FEM allows to reduce the order of derivatives by employing the test functions, which is referred as weakening. Once the weak form is obtained, it is discretized to construct a linear system of equations using trial functions for a set of basis functions. The order and type of the bases functions directly affect the quality and efficieny of solution convergence and accuracy.

In this work, NURBS are preferred for the discretization. It offers smooth crossing at element boundaries. Also its exact representation of round objects makes isoparametric formulation accessible for atomistic models, which cancels out geometric approximation errors. Compared to the Lagrange discretization, it achieves better convergence for both ground-state calculations and beyond [36].

The DFPT scheme within FEM requires recursive construction of the elements of linear system of equations within SCF iterations. These systems may be solved using either direct or iterative solvers with preconditioning or not. Moreover, it incorporates orthogonalization procedures to eliminate numerical instabilities in degenerate cases. For robust convergence, it benefits from mixing schemes on the density perturbations.

4.1 Weak formulation

Weak form formulation is constructed upon boundary conditions. In weakening of the Sternheimer equation (3.27), $\psi_\alpha^{(1)}$ is subject to the homogeneous Dirichlet boundary condition for $\psi_\alpha^{(0)}$ solved in the Kohn-Sham DFT with the same boundary condition. Introducing a test function φ, which theoretically can be any function in FEM and here is selected as $\delta\psi_\alpha^{(\mathbf{R}_I)}$, the Sternheimer equation takes the weak form of,

$$\begin{aligned}
-\langle\varphi(\hat{H}^{(0)}-\epsilon_\alpha^{(0)})\psi_\alpha^{(\mathbf{R}_I)}\rangle = &\int \nabla\varphi\lambda_I^{(0)}(\mathbf{r})\,d^3\mathbf{r}\int \lambda_I^{*(0)}(\mathbf{r}`)\psi_\alpha^{(0)}(\mathbf{r}`)\,d^3\mathbf{r}` \\
&+\int \varphi\lambda_I^{(0)}(\mathbf{r})\,d^3\mathbf{r}\int \lambda_I^{*(0)}(\mathbf{r}`)\nabla\psi_\alpha^{(0)}(\mathbf{r}`)\,d^3\mathbf{r}` \\
&+\int \varphi\left[\int \frac{\delta^2 E_{xc}[\rho^{(0)}]}{\delta\rho(\mathbf{r})^2}\rho'^{(\mathbf{R}_I)}\,d^3\mathbf{r}'\right]\psi_\alpha^{(0)}(\mathbf{r})\,d^3\mathbf{r} \\
&+\int \varphi\,[S_C]\,\psi_\alpha^{(0)}(\mathbf{r})\,d^3\mathbf{r},
\end{aligned} \tag{4.1}$$

where the zeroth-order kinetic energy contribution at the LHS is manipulated as $-\langle\varphi\nabla^2\psi_\alpha^{(\mathbf{R}_J)}\rangle = \langle\nabla\varphi\cdot\nabla\psi_\alpha^{(\mathbf{R}_J)}\rangle$ using the integration by parts, divergence theorem and properties of the homogeneous Dirichlet BC.

The discretization follows the weak form in FEM. The discrete forms of test and trial functions are given as $\sum_I v^I N_I$ for a set of basis functions N_I and solution DOFs v^I. The zeroth-order discrete Hamiltonian $[H^{(0)}]^{IJ}$ consists of kinetic energy, nonlocal potential, electrostatic, and exchange correlation terms,

$$\begin{aligned}
[H_T^{(0)}]^{IJ} &= \langle\nabla N_I\cdot\nabla N_J\rangle \\
[H_{NL}^{(0)}]^{IJ} &= \sum_{A=1}^{N}\sum_{k=1}^{k_A}\langle N_I\lambda_{Ak}^{(0)}\rangle h_{Ak}\langle\lambda_{Ak}^{(0)}N_J\rangle \\
[H_C^{(0)}]^{IJ} &= \frac{1}{2}\langle N_I v_c^{(0)}N_J\rangle \\
[H_{xc}^{(0)}]^{IJ} &= \langle N_I v_{xc}^{(0)}N_J\rangle.
\end{aligned} \tag{4.2}$$

Also the zeroth-order eigenvalue contributes to the LHS with the mass matrix $M^{IJ} = \langle N_I N_J\rangle$. On the other hand, the first-order discrete Hamiltonian $[H^{(1)}]^{IJ}$

contains nonlocal potential, electrostatic, and exchange correlation terms,

$$\begin{aligned}
[H_{NL}^{(1)}]^{IJ} &= \sum_{k=1}^{k_A} \left[\langle \nabla N_I \lambda_{Ak}^{(0)} \rangle h_{Ak} \langle \lambda_{Ak}^{(0)} N_J \rangle + \langle N_I \lambda_{Ak}^{(0)} \rangle h_{Ak} \langle \lambda_{Ak}^{(0)} \nabla N_J \rangle \right] \\
[H_C^{(1)}]^{IJ} &= \langle N_I v_c^{(1)} N_J \rangle \\
[H_{xc}^{(1)}]^{IJ} &= \langle N_I v_{xc}^{(1)} N_J \rangle.
\end{aligned} \tag{4.3}$$

Therefore, the discrete Sternheimer equation becomes,

$$\left([H^{(0)}]^{IJ} - \epsilon_\alpha^{(0)} [M]^{IJ} \right) \{\psi_\alpha^{(1)}\}^J = -[H^{(1)}]^{IJ} \{\psi_\alpha^{(0)}\}^J. \tag{4.4}$$

The Sternheimer equation also includes unperturbed and first-order perturbed electrostatic terms. In order to calculate these values, the Poisson problems (2.10) and equation (3.29) should be weakened and discretized. The weak forms are written as,

$$-\frac{1}{4\pi} < \varphi \nabla^2 \hat{v}_C > = < \varphi(\rho + b) >, \tag{4.5}$$

$$-\frac{1}{4\pi} < \varphi \nabla^2 S_C > = < \varphi(\rho^{(\mathbf{R}_I)} + b^{(\mathbf{R}_I)}) >, \tag{4.6}$$

where $S_C = \int \frac{\delta^2 E_C[\rho+b]}{\delta(\rho(\mathbf{r})+b(\mathbf{r}))^2} (\rho'^{(\mathbf{R}_J)} + b'^{(\mathbf{R}_J)})\, d^3\mathbf{r}'$. The LHS of equations (4.5) and (4.6) are prone to decrease the order of derivative by using integration by parts and divergence theorem. Therefore, the discretized electrostatic problems become,

$$\frac{1}{4\pi} \langle \nabla N_I \cdot \nabla N_J \rangle \{\hat{v}_C\}^J = \langle N_I (\rho + b) \rangle, \tag{4.7}$$

$$\frac{1}{4\pi} \langle \nabla N_I \cdot \nabla N_J \rangle \{S_C\}^J = \langle N_I (\rho^{(\mathbf{R}_I)} + b^{(\mathbf{R}_I)}) \rangle. \tag{4.8}$$

Weak form of the exact-discrete change contribution (3.42) is derived as,

$$\begin{aligned}
8 f_\alpha \langle \varphi \left(v_{ext}^{(0)} - \hat{v}_{ext}^{(0)} \right) \psi_\alpha^{(\mathbf{R}_I)} \rangle = &-4 f_\alpha \langle \varphi \left(v_{ext}^{(\mathbf{R}_I)} - \hat{v}_{ext}^{(\mathbf{R}_I)} \right) \psi_\alpha^{(0)} \rangle \\
&- 4 f_\alpha \langle \psi_\alpha^{(0)} \left(v_{ext}^{(\mathbf{R}_I)} - \hat{v}_{ext}^{(\mathbf{R}_I)} \right) \varphi \rangle.
\end{aligned} \tag{4.9}$$

4.2 NURBS discretization

Non-Uniform Rational B-Splines (NURBS) provide a flexible and compact basis for discretization capturing complex geometry within an isoparametric formulation [36]. In this work, NURBS basis functions [37] are employed for the numerical solutions of the Kohn-Sham and Sternheimer equations, and the Poisson problems arising in DFPT.

In three dimension, the numerical domain is defined in the parametric space $(\xi^1, \xi^2, \xi^3) \in [\xi^1_0, \xi^1_{m_1}] \times [\xi^2_0, \xi^2_{m_2}] \times [\xi^3_0, \xi^3_{m_3}]$, where each parametric direction $\alpha \in \{1, 2, 3\}$ is associated with a knot vector Ξ^α given as,

$$\Xi^\alpha = \{\xi^\alpha_0, ..., \xi^\alpha_{p^\alpha}, \xi^\alpha_{p^\alpha+1}, ..., \xi^\alpha_{n^\alpha}, ..., \xi^\alpha_{m^\alpha}\}. \tag{4.10}$$

Considering $m^\alpha = n^\alpha + p^\alpha + 1$, each knot vector Ξ^α contains $m^\alpha + 1$ non-decreasing knots ξ^α, which are non-uniformly distributed in general, such that the first and last $p^\alpha + 1$ of them have the same value with the open knot vectors.

The knot vector defines $n^\alpha + 1$ univariate B-spline basis functions $B^\alpha_a(\xi^\alpha)$, $a \in \{0, ..., n_\alpha\}$. These functions are non-negative, satisfy the partition of unity, have local support, and are piecewise polynomial of order p_α. Then, the trivariate NURBS basis functions are defined by,

$$R_{abc}(\xi^1, \xi^2, \xi^3) = \frac{w_{abc}}{W(\xi^1, \xi^2, \xi^3)} B^1_a(\xi^1) B^2_b(\xi^2) B^3_c(\xi^3), \tag{4.11}$$

with the positive normalizing function,

$$W(\xi^1, \xi^2, \xi^3) = \sum_{a=0}^{n_1} \sum_{b=0}^{n_2} \sum_{c=0}^{n_3} w_{abc} B^1_a(\xi^1) B^2_b(\xi^2) B^3_c(\xi^3), \tag{4.12}$$

for the weights $w_{abc} > 0$. Relying on this definition, the trivariate NURBS basis functions become non-negative, satisfy the partition of unity, have local support, and are piecewise polynomial rational.

Each patch is defined by the knot vectors in each parametric direction with the associated weights and the control points, which are the degrees of freedom (DOF) multiplying the basis functions. NURBS discretization guarantees interpolatory

behavior at the patch corners whereas it does not ensure interpolatory elsewhere unlike Lagrange discretization. Mathematically, the first $B_0^\alpha(\xi_0^\alpha) = 1$ and the last $B_{n^\alpha}^\alpha(\xi_{m^\alpha}^\alpha) = 1$ whereas $B_{a>0}^\alpha(\xi_0^\alpha) = 0$ and $B_{a<n^\alpha}^\alpha(\xi_{m^\alpha}^\alpha) = 0$. Therefore, the Dirichlet boundary conditions cannot be enforced on control points without weak enforcement strategies. A univariate B-spline of degree p^α is $C^{p^\alpha-1}$ continuous and each interior knot ξ_k^α repetition reduces the continuity order by one. For example, p^α times repeated knot results in C^0 continuity at that location by introducing a cusp in the basis function.

The discretized NURBS mesh of this work consists of 7 patches (see [36]). The central patch is a cube and each face connects to the remaining 6 patches, which wraps up to a sphere at the outer surface. Across these patches, only C^0 continuity is enforced by ensuring continuity at shared control points. The B-spline polynomial degree p^α is fixed in each parametric direction within each patch.

Let the knot vector $\bar{\Xi}^\alpha = \{\bar{\xi}_0^\alpha, ..., \bar{\xi}_{m^\alpha}^\alpha\}$ for a patch have a unique subset $\{\bar{\xi}_0^\alpha, ..., \bar{\xi}_{e^\alpha}^\alpha\}$ where e^α represents the number of unique knot spans. The element spans in the parametric space is defined through a tensor-product partitioning $e^1 \times e^2 \times e^3$. It prepares the NURBS mesh, similar to the Lagrange mesh, even though it does not yield a conventional mesh. The domain shape function are represented by N_I where I stands for the degrees of freedom. Therefore, the physical coordinates, i.e. $\mathbf{r}$ as the control points defining the geometry, and any scalar field, i.e. v as the degrees of freedom, are expressed as $\mathbf{r}(\xi) = \sum_I N_I(\xi)\mathbf{r}_I$ and $v(\xi) = \sum_I N_I(\xi)v_I$ in an isoparametric formulation, which formulates the domain geometry and solution fields with the same basis functions.

4.3 Solution scheme

The Sternheimer problem requires the knowledge of the zeroth-order wavefunctions, occupancy factors and eigenvalues for the Kohn-Sham problem. It is mandatory to have the Kohn-Sham DFT solutions beforehand. Similarly, the

Sternheimer equation searches approximate solutions to the true first-order wavefunctions, occupancy factors and eigenvalues minimizing the second-order Kohn-Sham energy perturbation which is proven to be variationally consistent.

The Sternheimer equations are solved via self-consistent field iterations [38]. The solution starts with a guess of the first-order wavefunction perturbation. After each iteration, a better first-order wavefunction perturbation is predicted. Before the SCF iterations, the LHS matrix is created and stored because of its first-order wavefunction independent structure. On the other hand, the RHS changes according to the preceding first-order electron density perturbations. The first-order electrostatic potential perturbations are also calculated via Poisson's problem for each density perturbation. Moreover, the first-order electron density perturbation depends on all the occupied states' wavefunctions for the specific perturbation. Therefore, each occupied states' perturbations are calculated at each iteration. For smoother convergence, the Anderson mixing scheme is implemented for the first-order electron density perturbation.

As the second-order energy perturbation has variational property with respect to the first-order wavefunction, the minimum wavefunction can be searched via both solvers for linear system of equations in SCF iterations and minimization algorithms. Both are prone to the preconditioners to speed up the convergence. In this work, iterative linear system solvers with preconditioner are preferred.

Both numerical and analytical aspects of the problem require the use of projectors at various stages. The first projector appearing on the RHS of the Sternheimer equation (3.46) ensures that the nonhomogeneous Sternheimer equation admits a solution by orthogonalizing the RHS to the homogeneous LHS solution. The second projector on the first-order wavefunction assures the orthogonality of the wavefunction to the Kohn-Sham solution. The third projector, the leftmost term of the LHS, is arbitrary for refined-meshes. However, $\hat{H}^{(0)} - \epsilon_\alpha^{(0)}$ cannot cancel out the zeroth-order Kohn-Sham orbital contributions for low resolution systems. Therefore, it is used to eliminate pollution in the zeroth-order wavefunction directions.

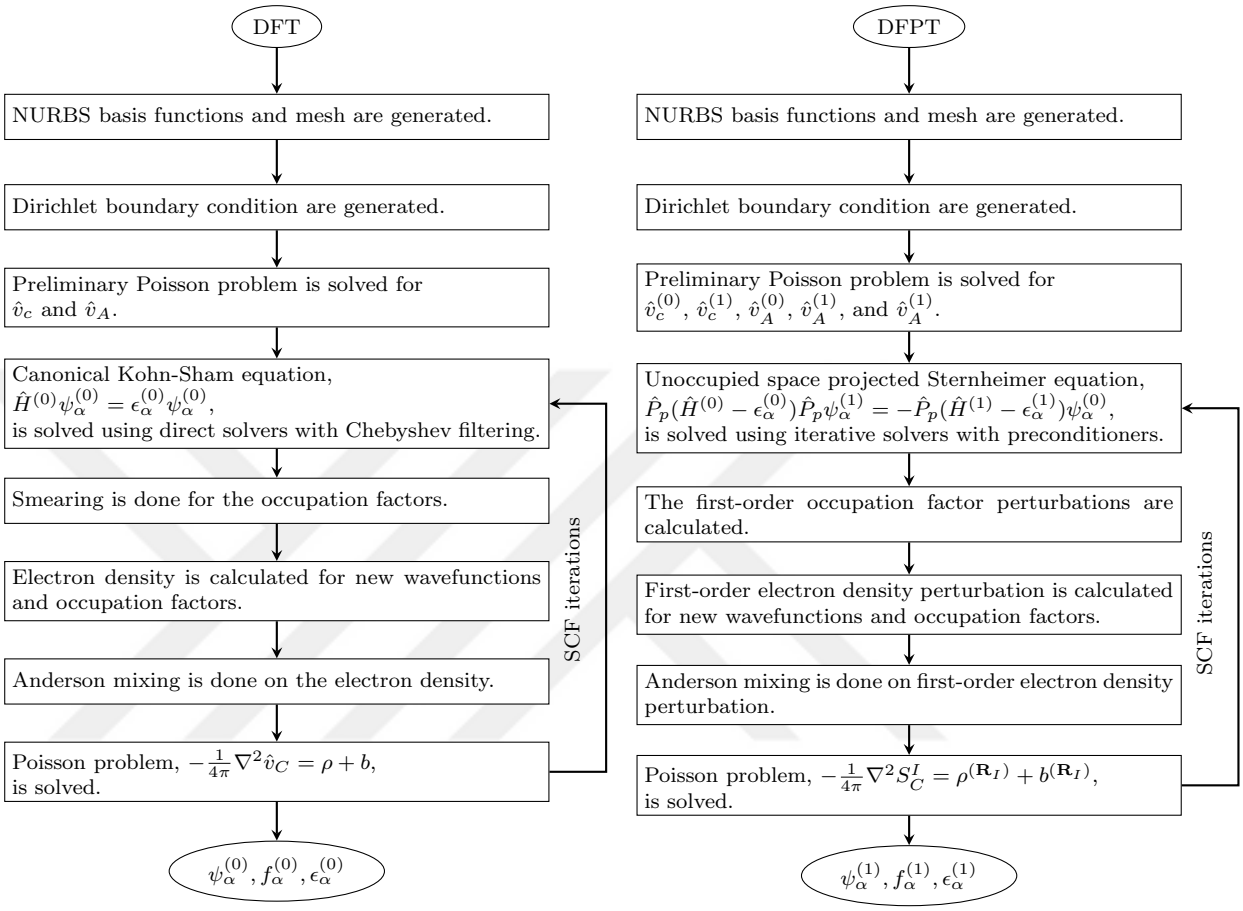

Figure 4.1: Kohn-Sham DFT and DFPT flowchart

The general frameworks of DFT and DFPT share structural similarities (see Figure 4.1). However, the Kohn-Sham DFT requires an eigenvalue problem solution with Chebyshev filtering whereas the DFPT involves to solve a singular linear system of equations by using iterative solvers. The inherent singularity in DFPT is addressed through projecting the linear system of equations and first-order wavefunction perturbations upon the unoccupied subspace, depending on the gauge selection. In comparison, the DFT enforces orthogonality of the zeroth-order wavefunctions through the Gram-Schmidt algorithm.

Chapter 5

Numerical Investigations

The Kohn-Sham DFT energy, forces, second-order perturbations and Sternheimer equations have been derived within the pseudopotential formalism. Then, the weak and discretized forms have been found in the FEM framework. In this chapter, the accuracy of DFPT derivations and implementations are confirmed.

In order to validate the DFPT implementation, two viable options are employed. The first one is comparison of the interatomic force constants with those obtained from contemporary codes such as ABINIT and Quantum Espresso. The latter is verification via FD using energy or analytical forces. Both approaches are used in different stages.

The current DFPT implementation in QAGE for the LDA exchange-correlation functional complies with ABINIT using the *Hartwingsen-Goedeker-Hutter (HGH)* pseudopotentials [23]. Therefore, DFPT may be tested with ABINIT for the LDA cases. On the other hand, the GGA implementation based on the *Gaussian pseudopotentials* of Willand is not supported in the contemporary codes. Therefore, the FD scheme has to be employed for the verified DFT energies and forces. In this study, the exchange-correlation functionals are accessed via Libxc [39].

The numerical investigations are separated to two main branches based on

the choice of LDA and GGA exchange-correlation functionals. In each brach, validation is carried out for molecules with categories of both non-degenerate and degenerate unperturbed orbital.

The first category consists of non-degenerate cases, where both the diagonal and parallel transport gauges are working. Any diatomic molecule in the first column of periodic table is suitable for this category. However, it is valid for the pseudopotential formalism because the first column molecules except dihydrogen become degenerate in the all electron setting due to the degeneracies in fully occupied core states. Also the remaining first column diatomics differ from dihydrogen by additional nonlocal contributions. Consequently, the dihydrogen (H_2) and dilithium (Li_2), and ammonia (NH_3) molecules may be analyzed. The second category involves degenerate cases. For this category, methane (CH_4) and dinitrogen (N_2) could be preferred.

All tests are performed using uniform C^2-continuous cubic B-spline discretization to generate the finite element mesh. To assess compatibility of pseudopotential calculations, *chemical accuracy*, defined as $1.6 \cdot 10^{-4}$ Hartree per atom, is targeted for absolute energy convergence. In order to demonstrate further accuracy, this threshold is exceeded in the subsequent examples. These examples display energy in Ha, force in Ha/a_0 and IFC in Ha/a_0^2 for the energy unit Hartree, Ha, and length unit Bohr, a_0.

Two axis convergence plots have the number of degree of freedom at the x-axis and error at the y-axis in logarithmic scale. As an instance, the energy error is defined as $E - E_0/M$, which is the difference between the current calculation and the reference energy. In order to comply with the chemical accuracy definition, the error is divided by the number of atoms in the system. Also the force and IFC errors has the same formulation.

5.1 Demonstration of LDA

The Libxc library allows to select exchange and correlation functionals separately. For the following examples, the Slater for exchange and the Wigner parametrization for correlation are chosen [40, 24].

5.1.1 Non-degenerate case

Dihydrogen molecule is modeled as two atoms in three-dimensional real space with coordinates given at Table 5.1 in Bohr, a_0. In the FEM basis, selection of patch dimensions requires a special care to observe the variational consistency and monotonous convergence. Specific to the dihydrogen LDA, central patch is selected as a square with dimensions of $3x3x3a_0^3$ whereas the radius of the outer patch is $30a_0$. For comparison, accuracy convergence up to $10^{-6}Ha$ in ABINIT is achieved with the *acell* of $30x30x30a_0^3$ and the *ecut* of $180Ha$. The *acell* refers to the lengths of primitive vectors whereas the *ecut* is the maximal plane-wave kinetic energy cut-off.

atomic positions in a_0

atom	x	y	z
H_1	-0.7400	0.0000	0.0000
H_2	0.7400	0.0000	0.0000

Table 5.1: Atomic positions for the dihydrogen with LDA

The reference values obtained from ABINIT are $E_0 = -1.1334114023Ha$ for the ground-state energy, $F_x^1 = 0.0045796579Ha/a_0$ for the force on the first atom in the x-direction and $C_{xx}^{11} = 0.30167Ha/a_0^2$ for the xx-component of the IFC tensor corresponding to the first atom acting on itself. For a fine mesh, QAGE delivers $E_0 = -1.1334117317Ha$, $F_x^1 = 0.0045799968Ha/a_0$ and $C_{xx}^{11} = 0.3016628085Ha/a_0^2$. Among these values, only the energy can be used as a reference for variational consistency. Comparing E_0' s reveals that the deviation

starts at the seventh digit for ABINIT and QAGE calculations converged up to $10^{-6} Ha$.

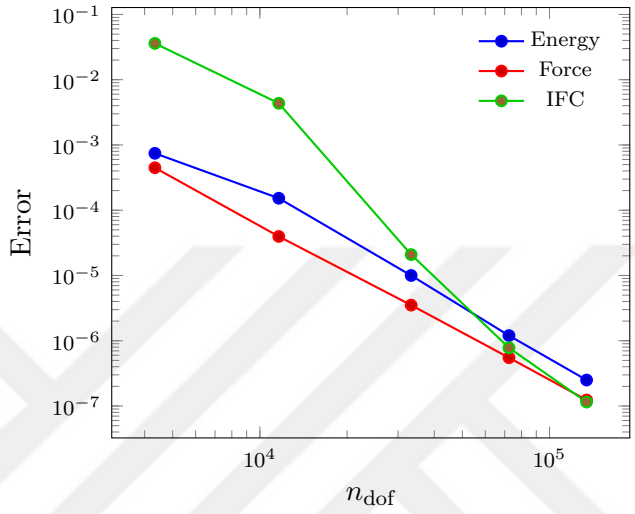

Figure 5.1: Dihydrogen convergence plot with the LDA exchange-correlation

The expression for the energy, force and second-order energy perturbation can be further analyzed through their *asymptotic convergence rate*. The asymptotic convergence rate for the energy is defined by,

$$E - E_0 = C n_{dof}^{-2k/3}, \tag{5.1}$$

where E_0 is the reference energy, n_{dof} is the number of degrees of freedom and C is a constant. In Figure 5.1, the slope of energy convergence in logarithmic scale, which is obtained by fitting selected data points, corresponds to $-2k/3$. The optimal asymptotic rate is equal to the order of discretization in FEM. This validation is legitimate only for energy and second-order energy perturbation but force is also inspected for a complementary perspective. The slope of energy, force and interatomic force constant are calculated for the last four discretization points. As a result, the asymptotic convergence rates are obtained as 4.03, 3.54, and 5.54. For a B^3 discretization, a rate over 3.00 is assessed as a super-optimal convergence rate for uniform B-spline.

Dilithium is a two atoms molecule situated in a three-dimensional space with atomic coordinates given in Table 5.2. The central patch is selected as a square

with dimensions of $3.75x3.75x3.75a_0^3$ and the outer patch radius is set to $30a_0$. On the other hand, ABINIT achieves energy convergence up to $10^{-7}Ha$ using *acell* of $80x80x80a_0^3$ and *ecut* of $25Ha$. This indicates that further refinement of these parameters changes the energy at most in the order of 10^{-8}.

atomic positions in a_0

atom	x	y	z
Li_1	-2.6000	0.0000	0.0000
Li_2	2.6000	0.0000	0.0000

Table 5.2: Atomic positions for the dilithium with LDA

The defined ABINIT system gives the reference values of $E_0 = -0.4459053841Ha$, $F_x^1 = 0.0016899499Ha/a_0$ and $C_{xx}^{11} = 0.01249Ha/a_0^2$. For a fine NURBS mesh, $E_0 = -0.4459076232Ha$, $F_x^1 = 0.0016914109Ha/a_0$ and $C_{xx}^{11} = 0.0124917857Ha/a_0^2$ are obtained from QAGE where the system energy also converged up to $10^{-7}Ha$. Even though the QAGE and ABINIT energies hold up to $10^{-5}Ha$, which satisfy the chemical accuracy, they are not the same up to the $10^{-7}Ha$ accuracy because of outer radius selection for the discretization patch. As an observation, it is critical to select a sufficiently large outer diameter to capture more accurate results in FEM. However, the limiting factor in this selection could be a slower convergence. In QAGE, the asymptotic convergence rates are obtained as 4.23, 5.07, and 5.28 calculating for the last four discretization points.

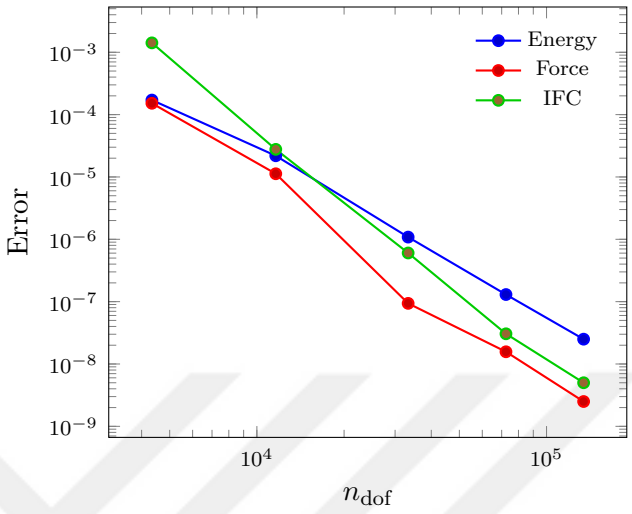

Figure 5.2: Dilithium convergence plot with the LDA exchange-correlation

5.1.2 Degenerate case

Degeneracy obliges the Sternheimer equation to be solved with the parallel transport gauge. Unless the parallel transport gauge is selected over the diagonal gauge, the solution DOFs diverge. As a degenerate case with LDA, methane molecule (CH_4) is inspected with the atomic positions given in Table 5.3. The main patch is set to a square with the dimensions of $3.00x3.00x3.00a_0^3$ and the outer patch radius is $110a_0$. As comparison, ABINIT converged up to $10^{-6}Ha$ at *acell* of $30x30x30a_0^3$ and *ecut* of $180Ha$.

atomic positions in a_0

atom	x	y	z
C_1	0.00000000	0.00000000	0.00000000
H_1	-1.20188799	-1.20188799	-1.20188799
H_2	1.20188799	1.20188799	-1.20188799
H_3	-1.20188799	1.20188799	1.20188799
H_4	1.20188799	-1.20188799	1.20188799

Table 5.3: Atomic positions for the methane with LDA

For the given specifications, ABINIT gives the reference values of E_0 =

$-8.0008720908Ha$, $F_x^1 = -0.0019601854Ha/a_0$ and $C_{xx}^{11} = 0.49831Ha/a_0^2$ while the QAGE values are $E_0 = -8.0008710901Ha$, $F_x^1 = -0.0019603894Ha/a_0$ and $C_{xx}^{11} = 0.4983116254Ha/a_0^2$ for a refined mesh. These energies agree in $10^{-6}Ha$ accuracy and this is also beyond chemical accuracy. The asymptotic convergence rates are found as 5.30, 6.38, and 6.65 calculating for the last four discretization points.

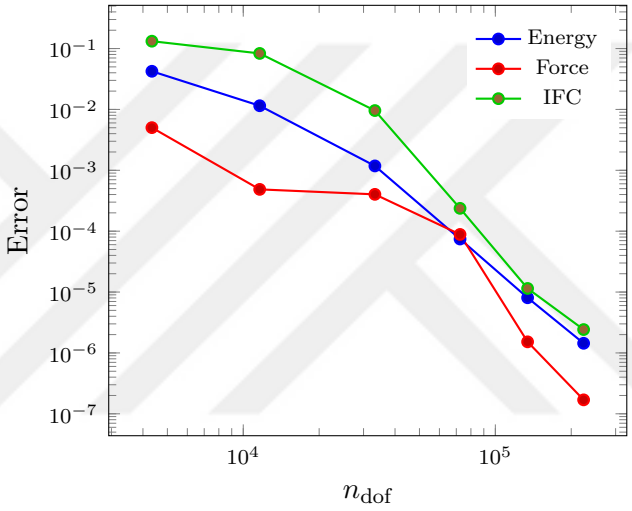

Figure 5.3: Methane convergence plot with the LDA exchange-correlation

5.2 Demonstration of GGA

For the GGA tests, the Perdew-Burke-Ernzerhof exchange-correlation functionals are used [4]. It also includes a core-correction to the electron density according to the Willand's Gaussian pseudopotentials. This correction affects the exchange-correlation functionals as well.

The chemical accuracy of energy and forces in GGA formulation had been proven in the previous studies [11]. Calculation of interatomic force constant tensors can be verified via central second-order derivatives for energy. The diagonal elements of interatomic force constant tensors can be derived using pure derivative formulations whereas the off-diagonal elements require mixed derivative

formulations. However, verification over energy is expensive. The fourth-order accurate FD of second-order derivative requires $4NS + 6NS(NS-1) + 1$ independent DFT energy calculations where N is the number of atoms and S is the problem dimension. Compared to the second-order formulations, the fourth-order accurate first-order derivative FD upon analytically taken first-order derivative drastically simplifies the number of necessary calculations. In DFT, the force calculation does not require any further endeavor beyond DFT SCF iterations and it is absolute within DFT. The central fourth-order accurate FD of first derivative needs only $4NS$ independent DFT energy calculations. Therefore, the latter option is invoked for verification in any DOF. Furthermore, it is significant to clarify that agreement of the IFC tensors calculated with FD and DFPT in any order suffices to prove the implementation accuracy due to variational structure of FEM.

5.2.1 Symmetry convergence

Given a system of N atoms in three-dimensional Cartesian coordinates, the IFC tensor is a real symmetric $3N$x$3N$ matrix. This tensor contains various symmetries originating from the translational and rotational invariance of the system as well as geometrical symmetries. As an instance of a two-atom system's IFC tensor, the first row's first column is equal in magnitude but opposite in sign to the first row's fourth column. Physically, the first is the response of the first atom to its own perturbation in the x-direction whereas the latter is the response of the first atom to the second atom's perturbation in the x-direction. Such symmetries offer stringent checks for the assessment of DFPT calculations.

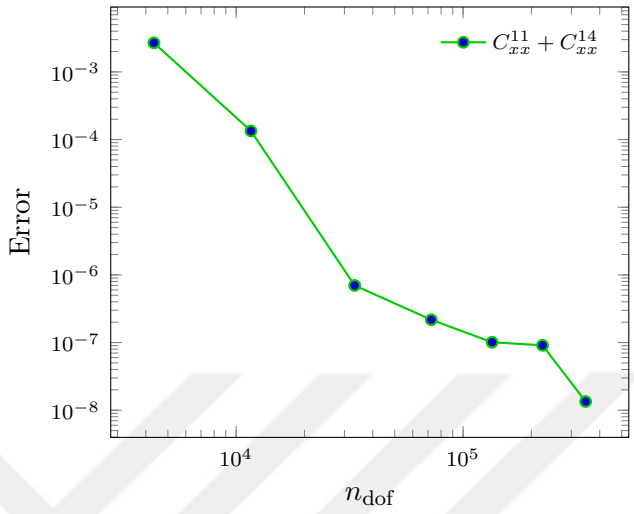

Figure 5.4: Dilithium symmetry convergence plot with the GGA exchange-correlation, depicting the absolute difference between the first row's first and fourth column elements of IFC

Even though the first row first and fourth column elements of the two atom system's IFC tensor physically possess the identity, they should numerically differ in any numerical framework, which already does not employ any symmetry identities to construct the tensor without separate calculations. However, these two values must converge to the same value as the mesh structure is refined. For a coarse mesh, the dilithium molecule's corresponding tensor values starts from $0.00564425 Ha/a_0^2$ and $-0.01101668 Ha/a_0^2$. As the mesh refined, these converge to $0.01054961 Ha/a_0^2$ and $-0.01054942 Ha/a_0^2$.

5.2.2 Mesh-independent coherence of DFPT and FD

During the LDA demonstration, the results are compared for converged QAGE and ABINIT energies. However, validation with FD may not need to achieve a full convergence in energy. In case of that the FD and DFPT IFCs agree even for a coarse mesh, it can reduce compute time necessary for the validation. Therefore, difference between the FD and DFPT IFCs' first-column first-row elements is observed with respect to the number of DOF for dilithium molecule with GGA.

Even though the aforementioned anomaly in the first-column first-row of IFC, which extremely deviates from the symmetric behavior, Figure 5.5 shows that compliance between DFPT and FD could be assumed independent of how refined the mesh is, in the aspect of chemical accuracy. Therefore, any order of mesh can be arbitrarily selected for validation.

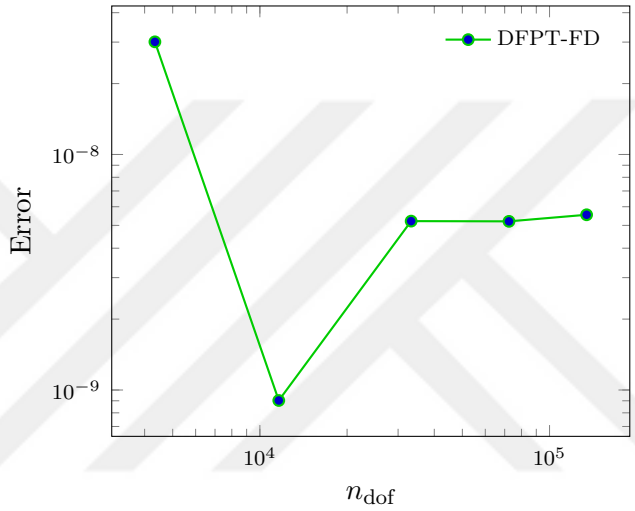

Figure 5.5: Investigation of FD and DFPT correlation with (Li_2)

5.2.3 Non-degenerate case

Dilithium is investigated with the GGA exchange-correlation functional as a non-degenerate example for the given coordinates in Table 5.4. The central square patch is switched to the dimensions of $4.50x4.50x4.50a_0^3$ and the outer patch radius is changed to $75a_0$.

atomic positions in a_0

atom	x	y	z
Li_1	-2.7000	0.0000	0.0000
Li_2	2.7000	0.0000	0.0000

Table 5.4: Atomic positions for the dilithium with GGA

For a fine mesh in this setting, the system energy, force, and first-row

first-column element of the IFC tensor converge to $E_0 = -0.7502969813 Ha$, $F_x^1 = 0.0032612054 Ha/a_0$ and $C_{xx}^{11} = 0.0105496054 Ha/a_0^2$, where the system energy converge up to $10^{-7} Ha$. Moreover, the asymptotic convergence rates are determined as 4.37 for energy, 3.63 for force, and 3.48 for IFC by calculating for the last five discretization points.

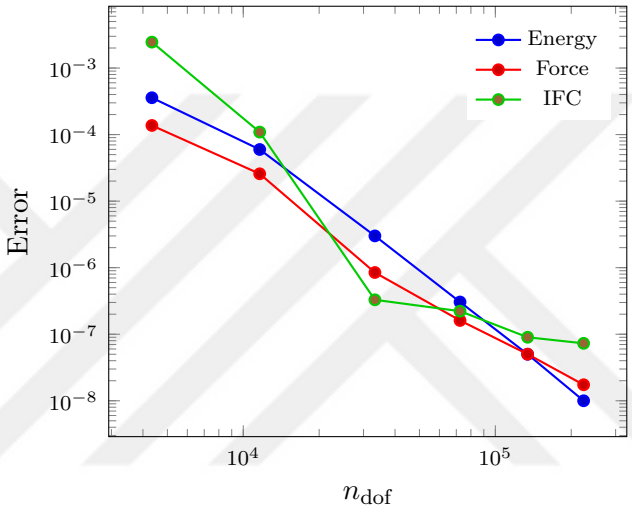

Figure 5.6: Dilithium convergence plot with the GGA exchange-correlation

The IFC tensors of dilithium via DFPT and FD are compared in a mesh with 11621 DOF. Major error among these components appear in the order of $10^{-8} Ha/a_0^2$.

$I\alpha$ $J\beta$	11	12	13	21	22	23
11	0.010331446	0.000000000	0.000000000	-0.010599412	0.000000000	0.000000000
12	0.000000000	0.001073257	0.000000000	0.000000000	-0.000614937	0.000000000
13	0.000000000	0.000000000	0.001073257	0.000000000	0.000000000	-0.000614937
21	-0.010599412	0.000000000	0.000000000	0.010331446	0.000000000	0.000000000
22	0.000000000	-0.000614937	0.000000000	0.000000000	0.001073257	0.000000000
23	0.000000000	0.000000000	-0.000614937	0.000000000	0.000000000	0.001073257

Table 5.5: IFC of dilithium with GGA in pseudopotential setting by DFPT

$I\alpha$ $J\beta$	11	12	13	21	22	23
11	0.010331448	0.000000000	0.000000000	-0.010599421	0.000000000	0.000000000
12	0.000000000	0.001073254	0.000000000	0.000000000	-0.000614947	0.000000000
13	0.000000000	0.000000000	0.001073254	0.000000000	0.000000000	-0.000614947
21	-0.010599421	0.000000000	0.000000000	0.010331448	0.000000000	0.000000000
22	0.000000000	-0.000614947	0.000000000	0.000000000	0.001073254	0.000000000
23	0.000000000	0.000000000	-0.000614947	0.000000000	0.000000000	0.001073254

Table 5.6: IFC of dilithium with GGA in pseudopotential setting by FD

Ammonia (NH_3) is also selected for its behavior because it does not display geometrical symmetries for the given coordinates in Table 5.7. This ammonia configuration is non-degenerate for the Perdew-Burke-Ernzerhof exchange-correlation functionals with the Willand's Gaussian pseudopotentials. For this case, the square patch has the dimensions of $3.00x3.00x3.00a_0^3$ and the outer patch radius is selected as $90a_0$.

atomic positions in a_0

atom	x	y	z
N_1	0.0000	0.0000	0.0000
H_1	0.0000	-0.0100	1.0280
H_2	0.8410	-0.4890	0.2940
H_3	-0.8410	-0.4890	0.2940

Table 5.7: Atomic positions for the ammonia with GGA

In Figure 5.7, force on the nitrogen in the y-direction is selected different than the previous selections. The reference values via QAGE are determined as $E_0 = -12.3146515366Ha$, $F_y^1 = 0.0080691848Ha/a_0$ and $C_{xx}^{11} = 0.6633141630Ha/a_0^2$. In this configuration, energy convergence up to $10^{-5}Ha$ is reached. For the assessment of an accurate implementation, the asymptotic convergence rates give 4.71 for energy, 4.27 for force, and 6.88 for IFC by calculating for the last five discretization points.

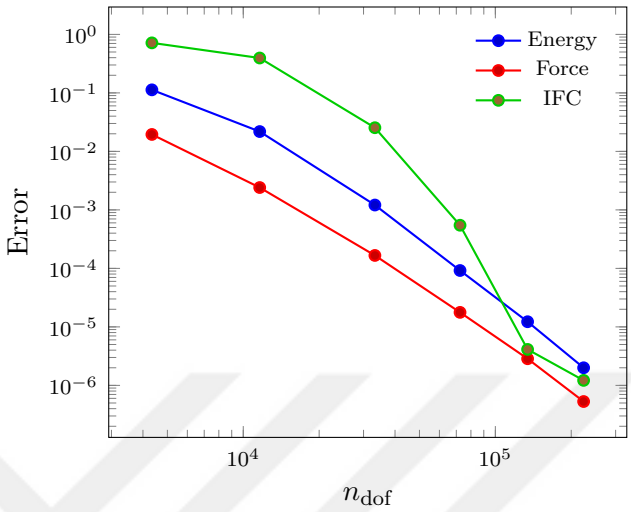

Figure 5.7: Ammonia convergence plot with the GGA exchange-correlation

The IFC tensor' s only interactions relevant to the nitrogen and the second hydrogen atoms are included in Tables 5.8 and 5.9. The biggest error in the given part of ammonia IFC is in the order of $10^{-6} Ha/a_0^2$ for a mesh corresponding to 11621 DOF.

$I\alpha$ $J\beta$	11	12	13	31	32	33
11	-0.915790663	0.000000000	0.000000000	-0.301163527	0.184533294	0.066270875
12	0.000000000	-2.059228684	0.149491866	0.139307502	-0.111966328	-0.057877968
13	0.000000000	0.149491866	-1.123363570	0.093057453	-0.090268017	-0.068328091
31	-0.301163527	0.139307502	0.093057453	0.404651784	-0.178901424	-0.092720576
32	0.184533294	-0.111966328	-0.090268017	-0.178901424	0.191431137	0.077540990
33	0.066270875	-0.057877968	-0.068328091	-0.092720576	0.077540990	0.202074581

Table 5.8: IFC of ammonia with GGA in pseudopotential setting by DFPT

$I\alpha$ $J\beta$	11	12	13	31	32	33
11	-0.915793074	-0.000000003	-0.000000001	-0.301164007	0.184533292	0.066270874
12	-0.000000003	-2.059232249	0.149492235	0.139307500	-0.111967038	-0.057877893
13	-0.000000001	0.149492235	-1.123365940	0.093057452	-0.090267944	-0.068328565
31	-0.301164007	0.139307500	0.093057452	0.404651941	-0.178901284	-0.092720599
32	0.184533292	-0.111967038	-0.090267944	-0.178901284	0.191431298	0.077540908
33	0.066270874	-0.057877893	-0.068328565	-0.092720599	0.077540908	0.202074815

Table 5.9: IFC of ammonia with GGA in pseudopotential setting by FD

5.2.4 Degenerate case

For the coordinates given in Figure 5.10, dinitrogen molecule (N_2) is investigated as a degenerate case. The central square patch for this system is chosen as $2.50x2.50x2.50a_0^3$ and the outer patch radius is selected as $90a_0$.

atomic positions in a_0

atom	x	y	z
N_1	-1.1350	0.0000	0.0000
N_2	1.1350	0.0000	0.0000

Table 5.10: Atomic positions for the dinitrogen with GGA

This system converges to the reference values of $E_0 = -21.0353204649Ha$, $F_y^1 = 0.1914899620Ha/a_0$ and $C_{xx}^{11} = 0.6788092190Ha/a_0^2$. For this system, the energy convergence up to $10^{-5}Ha$ is obtained. After the fitting, the asymptotic convergence rates give 4.29 for energy, 5.14 for force, and 4.85 for IFC by calculating for the last five discretization points.

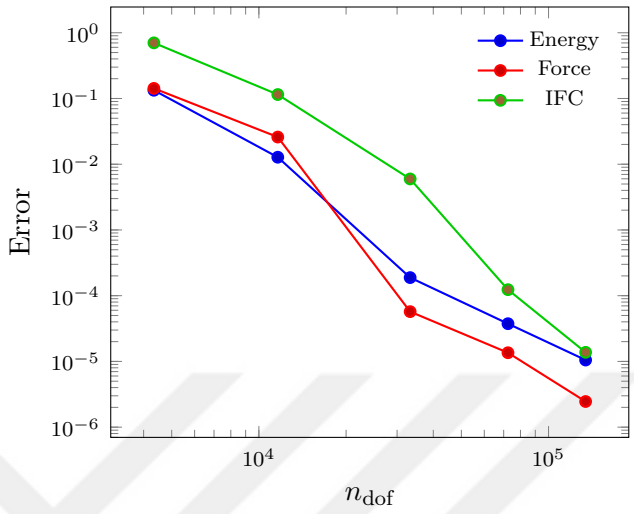

Figure 5.8: Dinitrogen convergence plot with the GGA exchange-correlation

Dominating error in the IFC tensor of dinitrogen is in the order of $10^{-7} Ha/a_0^2$ for a mesh corresponding to 11621 DOF. It achieves more than the chemical accuracy with monotonic convergence on each component.

$I\alpha$ $J\beta$	11	12	13	21	22	23
11	0.908436406	0.000000000	0.000000000	-0.624605516	0.000000000	0.000000000
12	0.000000000	-0.450722540	0.000000000	0.000000000	-0.087986756	0.000000000
13	0.000000000	0.000000000	-0.450722540	0.000000000	0.000000000	-0.087986756
21	-0.624605516	0.000000000	0.000000000	0.908436406	0.000000000	0.000000000
22	0.000000000	-0.087986756	0.000000000	0.000000000	-0.450722540	0.000000000
23	0.000000000	0.000000000	-0.087986756	0.000000000	0.000000000	-0.450722540

Table 5.11: IFC of dinitrogen with GGA in pseudopotential setting by DFPT

$I\alpha$ $J\beta$	11	12	13	21	22	23
11	0.908435988	0.000000000	0.000000000	-0.624605875	0.000000000	0.000000000
12	0.000000000	-0.450722645	0.000000000	0.000000000	-0.087986864	0.000000000
13	0.000000000	0.000000000	-0.450722645	0.000000000	0.000000000	-0.087986864
21	-0.624605875	0.000000000	0.000000000	0.908435988	0.000000000	0.000000000
22	0.000000000	-0.087986864	0.000000000	0.000000000	-0.450722645	0.000000000
23	0.000000000	0.000000000	-0.087986864	0.000000000	0.000000000	-0.450722645

Table 5.12: IFC of dinitrogen with GGA in pseudopotential setting by FD

Chapter 6

Conclusion

DFPT offers a DFT-based scheme systematically deriving higher-order material properties through perturbations in atomic positions, electric fields, and strains. In this work, DFPT formulation has been developed within the FEM framework for second-order nuclear perturbations. Compared to the planewave implementations, which is in reciprocal space, the FEM' s real-space basis provides greater flexibility in handling aperiodic systems.

The pseudopotential formulation has been validated against both popular DFPT codes and internal FD procedures. The LDA exchange-correlation functionals with pseudopotentials' results match ABINIT calculations beyond chemical accuracy while the GGA with pseudopotentials implementation achieves strong agreement between DFPT and FD IFC tensors. Moreover, the all-electron DFPT' s IFCs computed with LDA complies with the FD' s IFC, representing the first reported example of all-electron DFPT calculations in real space (see Appendix A).

The DFPT implementation in QAGE benefits from CPU parallelization at all available stages. However, given an efficient linear system solver, the FEM element calculations and assembly become the primary bottleneck. Porting these from CPU to GPU parallelization could significantly reduce compute time.

The singularity in Sternheimer equation necessitates either iterative solvers or direct solvers with pivoting. While available preconditioners accelerate convergence for basic molecular structures, they result in long computation times for complex Hamiltonian patterns. Developing a preconditioner tailored specifically to the Sternheimer equation could contribute to convergence while it also contributes to the computational materials community.

Implementation of the DFPT core enables both vertical and horizontal theoretical extensions. Vertically, once a framework is created for a specific perturbation, higher-order perturbations can be trivially obtained. Horizontally, inclusion of strain and electric field perturbations causes more initial costs, particularly due to the consideration of how to impose these perturbations. Nevertheless, the horizontal expansion offers valuable material properties, including the elastic tensor (second-order strain perturbation of energy) and the force-response internal-strain tensor (mixed perturbation of strain and atomic positions of energy). Ultimately, predicting nonlinear material behavior could be a greater goal.

Bibliography

[1] W. Kohn and L. J. Sham, "Self-consistent equations including exchange and correlation effects," *Physical Review*, vol. 140, pp. A1133–A1138, 1965.

[2] Hamann, Schluter, and Chiang, "Norm-conserving pseudopotentials," *Physical Review Letters*, vol. 43, pp. 1494–1497, 1979.

[3] J. P. Perdew and A. Zunger, "Self-interaction correction to density-functional approximations for many-electron systems," *Physical Review B*, vol. 23, pp. 5048–5079, 1981.

[4] J. P. Perdew, K. Burke, and M. Ernzerhof, "Generalized gradient approximation made simple," *Physical Review Letters*, vol. 77, pp. 3865–3868, 1996.

[5] J. B. Foresman, M. Head-Gordon, J. A. Pople, M. J. Frisch, C. Gonzalez, and J. S. Stephens, "Efficient implementation of density functional theory for polyatomic molecules," *Chemical Physics Letters*, vol. 239, no. 5-6, pp. 497–506, 1995.

[6] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, A. D. Corso, S. de Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A. P. Seitsonen, A. Smogunov, P. Umari, and R. M. Wentzcovitch, "Quantum espresso: a modular and open-source software project for quantum simulations of materials," *Journal of Physics: Condensed Matter*, vol. 21, no. 39, p. 395502, 2009.

[7] X. Gonze, J.-M. Beuken, R. Caracas, F. Detraux, M. Fuchs, G.-M. Rignanese, L. Sindic, M. Verstraete, G. Zerah, F. Jollet, M. Torrent, A. Roy, M. Mikami, P. Ghosez, J.-Y. Raty, and D. C. Allan, "First-principles computation of material properties: the abinit software project," *Computational Materials Science*, vol. 25, no. 3, pp. 478–492, 2002.

[8] P. F. Batcho and G. K. Schenter, "Computational method for general multicenter electronic structure calculations," *Physical Review E*, vol. 61, no. 6, pp. 7169–7179, 2000.

[9] S. Ghosh and P. Suryanarayana, "Sparc: Accurate and efficient finite-difference formulation and parallel implementation of density functional theory: Extended systems," *Computer Physics Communications*, vol. 216, pp. 109–125, 2017.

[10] P. Suryanarayana, V. Gavini, T. Blesgen, K. Bhattacharya, and M. Ortiz, "Non-periodic finite-element formulation of kohn–sham density functional theory," *Journal of the Mechanics and Physics of Solids*, vol. 58, no. 2, pp. 256–280, 2010.

[11] K. Karaca and Temizer, "Variationally consistent hellmann–feynman forces in the finite element formulation of kohn–sham density functional theory," *Computer Methods in Applied Mechanics and Engineering*, vol. 403, p. 115674, 2023.

[12] Temizer, "On the identification and finite element treatment of macroscopic stress in kohn–sham density functional theory," *Computer Methods in Applied Mechanics and Engineering*, vol. 435, p. 117629, 2025.

[13] B. Kanungo, P. M. Zimmerman, and V. Gavini, "Exact exchange-correlation potentials from ground-state electron densities," *Nature Communications*, vol. 10, no. 1, p. 4497, 2019.

[14] A. Sharma and P. Suryanarayana, "Calculation of phonons in real-space density functional theory," *Phys. Rev. E*, vol. 108, p. 045302, Oct 2023.

[15] S. Baroni, S. de Gironcoli, A. Dal Corso, and P. Giannozzi, "Phonons and related crystal properties from density-functional perturbation theory," *Rev. Mod. Phys.*, vol. 73, pp. 515–562, Jul 2001.

[16] X. Gonze, "Perturbation expansion of variational principles at arbitrary order," *Phys. Rev. A*, vol. 52, pp. 1086–1095, Aug 1995.

[17] E. A. Hylleraas, "Neue berechnung der energie des heliums im grundzustande, sowie des tiefsten terms von ortho-helium," *Zeitschrift für Physik*, vol. 65, pp. 209–225, 1930.

[18] O. Sinanoğlu, "Relation of perturbation theory to variation method," *The Journal of Chemical Physics*, vol. 34, pp. 1237–1240, 04 1961.

[19] A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. Mineola, New York: Dover Publications, 1996.

[20] R. G. Parr and W. Yang, *Density-Functional Theory of Atoms and Molecules*. New York: Oxford University Press, 1989.

[21] R. M. Martin, *Electronic Structure: Basic Theory and Practical Methods*. Cambridge: Cambridge University Press, 2004.

[22] A. Willand, Y. O. Kvashnin, L. Genovese, Álvaro Vázquez-Mayagoitia, A. K. Deb, A. Sadeghi, T. Deutsch, and S. Goedecker, "Norm-conserving pseudopotentials with chemical accuracy compared to all-electron calculations," *The Journal of Chemical Physics*, vol. 138, no. 10, p. 104109, 2013.

[23] C. Hartwigsen, S. Goedecker, and J. Hutter, "Relativistic separable dual-space gaussian pseudopotentials from h to rn," *Phys. Rev. B*, vol. 58, pp. 3641–3662, Aug 1998.

[24] E. Wigner, "On the interaction of electrons in metals," *Phys. Rev.*, vol. 46, pp. 1002–1011, Dec 1934.

[25] S. Lehtola, C. Steigemann, M. J. T. Oliveira, and M. A. L. Marques, "Recent developments in libxc - a comprehensive library of functionals for density functional theory," *SoftwareX*, vol. 7, pp. 1–5, 2018.

[26] P. Motamarri, M. Nowak, K. Leiter, J. Knap, and V. Gavini, "Higher-order adaptive finite-element methods for kohn–sham density functional theory," *Journal of Computational Physics*, vol. 253, pp. 308–343, 2013.

[27] J. E. Pask and P. A. Sterne, "Real-space formulation of the electrostatic potential and total energy of solids," *Phys. Rev. B*, vol. 71, p. 113101, Mar 2005.

[28] R. P. Feynman, "Forces in molecules," *Phys. Rev.*, vol. 56, pp. 340–343, Aug 1939.

[29] S. Baroni, P. Giannozzi, and A. Testa, "Green's-function approach to linear response in solids," *Phys. Rev. Lett.*, vol. 58, pp. 1861–1864, May 1987.

[30] X. Wu, D. Vanderbilt, and D. R. Hamann, "Systematic treatment of displacements, strains, and electric fields in density-functional perturbation theory," *Phys. Rev. B*, vol. 72, p. 035105, Jul 2005.

[31] R. M. Sternheimer, "Electronic polarizabilities of ions from the hartree-fock wave functions," *Phys. Rev.*, vol. 96, pp. 951–968, Nov 1954.

[32] X. Gonze and J.-P. Vigneron, "Density-functional approach to nonlinear-response coefficients of solids," *Phys. Rev. B*, vol. 39, pp. 13120–13128, Jun 1989.

[33] J. J. Sakurai and J. Napolitano, *Modern Quantum Mechanics*. Boston: Addison-Wesley, 2nd ed., 2011.

[34] S. T. Epstein, "Constraints and the v^{2n+1} theorem," *Chemical Physics Letters*, vol. 70, pp. 201–204, Mar 1980.

[35] X. Gonze, "Adiabatic density-functional perturbation theory," *Phys. Rev. A*, vol. 52, pp. 1096–1114, Aug 1995.

[36] Temizer, P. Motamarri, and V. Gavini, "Nurbs-based non-periodic finite element framework for kohn-sham density functional theory calculations," *Journal of Computational Physics*, vol. 410, p. 109364, 2020.

[37] L. Piegl and W. Tiller, *The NURBS Book*. Berlin: Springer, 2nd ed., 1996.

[38] 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.*, vol. 64, pp. 1045–1097, Oct 1992.

[39] M. A. Marques, M. J. Oliveira, and T. Burnus, "Libxc: A library of exchange and correlation functionals for density functional theory," *Computer Physics Communications*, vol. 183, p. 2272–2281, Oct. 2012.

[40] J. C. Slater and J. C. Phillips, "Quantum theory of molecules and solids vol. 4: The self-consistent field for molecules and solids," *Physics Today*, vol. 27, pp. 49–50, 12 1974.

[41] E. B. Wilson, J. C. Decius, and P. C. Cross, *Molecular Vibrations: The Theory of Infrared and Raman Vibrational Spectra.* New York: McGraw-Hill, 1955.

Appendix A

All-Electron

All-electron (AE) DFT calculations consider both core and valence electrons. Even though it is more intuitive than the pseudopotentials, it increases the number of orbitals to calculate. Moreover, convergence in calculations necessitates too fine meshes. Therefore, it is not preferable yet it is going to be detailed for completeness.

In the AE framework, each nuclei is formulated with the point-charge expression of $\beta_I = -Z_I\delta(\mathbf{r} - \mathbf{R}_I)$ with Z_I as the nuclear charge, where the nuclear potential takes the form of $v_I = -Z_I/r_I$. The AE energy definition differs from pseudopotential energy definition (2.2) by the dropped nonlocal potential contribution and the nuclear charges. Also the nonlinear core-correction to the electron density becomes irrelevant. Then, the AE energy definition with exact-discrete change is defined as,

$$E = 2\sum_{\alpha}^{occ/2} f_\alpha \langle \psi_\alpha^* \hat{H} \psi_\alpha \rangle - \langle \frac{\delta E_C}{\delta(\rho + b)} \rho \rangle + E_C - \langle \frac{\delta E_{xc}}{\delta \rho} \rho \rangle + E_{xc} - \frac{1}{2}\langle b\hat{v}_{ext} \rangle + E_{II}^{\delta}. \tag{A.1}$$

Therefore, the canonical Kohn-Sham equation equivalent to (2.17) becomes,

$$(-\frac{1}{2}\nabla^2 + v_c + v_{xc})\psi_\alpha = \epsilon_\alpha \psi_\alpha. \tag{A.2}$$

The Poisson problem for the electrostatic potential does not change except the ball-charge and point-charge at the RHS of (2.10). Weakened and discretized versions of (A.2) can be found following the same procedures in Chapter 4.

The second-order atomic positions perturbation of the AE energy, corresponding to (3.26), takes the form of,

$$\begin{aligned}
\frac{d^2E}{d\mathbf{R}_I d\mathbf{R}_J} &= 4\sum_{\alpha=1}^{occ/2} f_\alpha \langle \psi_\alpha^{(\mathbf{R}_I)} (\hat{H}^{(0)} - \epsilon_\alpha^{(0)}) \psi_\alpha^{(\mathbf{R}_J)} \rangle \\
&+ \int \left[\int \frac{\delta^2 E_{xc}[\rho^{(0)}]}{\delta\rho(\mathbf{r})^2} \rho'^{(\mathbf{R}_I)} \delta(\mathbf{r}' - \mathbf{r})\, d^3\mathbf{r}' \right] \rho^{(\mathbf{R}_J)}\, d^3\mathbf{r} \\
&+ \int \left[\int \frac{\delta^2 E_C[\rho + b]}{\delta(\rho(\mathbf{r}) + b(\mathbf{r}))^2} (\rho'^{(\mathbf{R}_J)} + b'^{(\mathbf{R}_J)})\, d^3\mathbf{r}' \right] (\rho^{(\mathbf{R}_I)} + b^{(\mathbf{R}_I)})\, d^3\mathbf{r} \\
&+ \int \frac{\delta E_C[\rho + b]}{\delta(\rho(\mathbf{r}) + b(\mathbf{r}))} b^{(\mathbf{R}_I, \mathbf{R}_I)} \delta_{IJ}\, d^3\mathbf{r} \\
&- \frac{1}{2} \int \hat{v}_I[\beta_I]^{(\mathbf{R}_I, \mathbf{R}_I)} b^{(0)} \delta_{IJ}\, d^3\mathbf{r} - \frac{1}{2} \int \hat{v}_I[\beta_I]^{(\mathbf{R}_I)} \beta_J^{(\mathbf{R}_J)}\, d^3\mathbf{r} \\
&- \frac{1}{2} \int \hat{v}_J[\beta_J]^{(\mathbf{R}_J)} \beta_I^{(\mathbf{R}_I)}\, d^3\mathbf{r} - \frac{1}{2} \int \hat{v}_{ext} \beta_I^{(\mathbf{R}_I, \mathbf{R}_I)} \delta_{IJ}\, d^3\mathbf{r} \\
&+ E_{II}^{\delta(\mathbf{R}_I, \mathbf{R}_J)}.
\end{aligned} \tag{A.3}$$

Derivative of point-charges diverges near the core so that manipulating the perturbation upon the point-charges could increase accuracy, which appear from the third to the sixth lines in (A.3). It can be achieved with the integration by parts and the gradient theorem by considering that $-\partial/\partial\mathbf{r}$ is equivalent to $\partial/\partial\mathbf{R}_I$ when acting on $\mathbf{r}_I$. It also requires to choose a big enough domain to eliminate the surface terms. For $S_C^I = \int \frac{\delta^2 E_C[\rho+b]}{\delta(\rho(\mathbf{r})+b(\mathbf{r}))^2} (\rho'^{(\mathbf{R}_I)} + b'^{(\mathbf{R}_I)})\, d^3\mathbf{r}'$, (A.3) can be revised

as,

$$\begin{aligned}\frac{d^2E}{d\mathbf{R}_I d\mathbf{R}_J} &= 4\sum_{\alpha=1}^{occ/2} f_\alpha \langle \psi_\alpha^{(\mathbf{R}_I)} (\hat{H}^{(0)} - \epsilon_\alpha^{(0)}) \psi_\alpha^{(\mathbf{R}_J)} \rangle \\ &+ \int \left[\int \frac{\delta^2 E_{xc}[\rho^{(0)}]}{\delta\rho(\mathbf{r})^2} \rho'^{(\mathbf{R}_I)} \delta(\mathbf{r}' - \mathbf{r})\, d^3\mathbf{r}' \right] \rho^{(\mathbf{R}_J)}\, d^3\mathbf{r} \\ &+ \int S_C^J \rho^{(\mathbf{R}_I)}\, d^3\mathbf{r} + \int \nabla S_C^J \beta_I\, d^3\mathbf{r} \\ &+ \int \nabla^2 \left(\frac{\delta E_C[\rho + b]}{\delta(\rho(\mathbf{r}) + b(\mathbf{r}))} \right) \beta_I^{(0)} \delta_{IJ}\, d^3\mathbf{r} \\ &- \frac{1}{2}\int \hat{v}_I[\beta_I]^{(\mathbf{R}_I, \mathbf{R}_I)} b^{(0)} \delta_{IJ}\, d^3\mathbf{r} - \frac{1}{2}\int \nabla \hat{v}_I[\beta_I]^{(\mathbf{R}_I)} \beta_J^{(0)}\, d^3\mathbf{r} \\ &- \frac{1}{2}\int \nabla \hat{v}_J[\beta_J]^{(\mathbf{R}_J)} \beta_I^{(0)}\, d^3\mathbf{r} - \frac{1}{2}\int \nabla^2 \hat{v}_{ext} \beta_I^{(0)} \delta_{IJ}\, d^3\mathbf{r} \\ &+ E_{II}^{\delta(\mathbf{R}_I, \mathbf{R}_J)}.\end{aligned} \tag{A.4}$$

Then, the Sternheimer equation for the AE energy, corresponding to (3.27), takes the form of,

$$\begin{aligned}8f_\alpha \langle \delta\psi_\alpha^{(\mathbf{R}_I)} (\hat{H}^{(0)} - \epsilon_\alpha^{(0)}) \psi_\alpha^{(\mathbf{R}_I)} \rangle &= -8f_\alpha \int \delta\psi_\alpha^{(\mathbf{R}_I)} \left[\int \frac{\delta^2 E_{xc}[\rho^{(0)}]}{\delta\rho(\mathbf{r})^2} \rho'^{(\mathbf{R}_J)}\, d^3\mathbf{r}' \right] \psi_\alpha^{(0)}(\mathbf{r})\, d^3\mathbf{r} \\ &- 8f_\alpha \int \delta\psi_\alpha^{(\mathbf{R}_I)} S_C^I \psi_\alpha^{(0)}(\mathbf{r})\, d^3\mathbf{r}.\end{aligned} \tag{A.5}$$

Similarly, the first-order perturbation of the Poisson equation for the electrostatic term is given as,

$$-\frac{1}{4\pi}\nabla^2 \left(\int \frac{\delta^2 E_C[\rho + b]}{\delta(\rho(\mathbf{r}) + b(\mathbf{r}))^2} (\rho'^{(\mathbf{R}_J)} + b'^{(\mathbf{R}_J)})\, d^3\mathbf{r}' \right) = \rho^{(\mathbf{R}_I)} + b^{(\mathbf{R}_I)}. \tag{A.6}$$

After the test function selection of φ as $\delta\psi_\alpha^{(\mathbf{R}_I)}$, the weakened form of (A.5) becomes,

$$\begin{aligned}-\langle \varphi (\hat{H}^{(0)} - \epsilon_\alpha^{(0)}) \psi_\alpha^{(\mathbf{R}_I)} \rangle &= \int \varphi \left[\int \frac{\delta^2 E_{xc}[\rho^{(0)}]}{\delta\rho(\mathbf{r})^2} \rho'^{(\mathbf{R}_I)}\, d^3\mathbf{r}' \right] \psi_\alpha^{(0)}(\mathbf{r})\, d^3\mathbf{r} \\ &+ \int \varphi \left[\int \frac{\delta^2 E_C[\rho + b]}{\delta(\rho(\mathbf{r}) + b(\mathbf{r}))^2} (\rho'^{(\mathbf{R}_I)} + b'^{(\mathbf{R}_I)})\, d^3\mathbf{r}' \right] \psi_\alpha^{(0)}(\mathbf{r})\, d^3\mathbf{r}.\end{aligned} \tag{A.7}$$

After the LHS's discrete matrices are determined as,

$$\begin{aligned}
[H_T^{(0)}]^{IJ} &= \langle \nabla N_I \cdot \nabla N_J \rangle \\
[H_C^{(0)}]^{IJ} &= \frac{1}{2} \langle N_I v_c^{(0)} N_J \rangle \\
[H_{xc}^{(0)}]^{IJ} &= \langle N_I v_{xc}^{(0)} N_J \rangle,
\end{aligned} \tag{A.8}$$

and the RHS's discrete matrices are determined as,

$$\begin{aligned}
[H_C^{(1)}]^{IJ} &= \langle N_I v_c^{(1)} N_J \rangle \\
[H_{xc}^{(1)}]^{IJ} &= \langle N_I v_{xc}^{(1)} N_J \rangle,
\end{aligned} \tag{A.9}$$

the discrete Sternheimer equation becomes,

$$\left([H^{(0)}]^{IJ} - \epsilon_\alpha^{(0)} [M]^{IJ} \right) \{\psi_\alpha^{(1)}\}^J = -[H^{(1)}]^{IJ} \{\psi_\alpha^{(0)}\}^J. \tag{A.10}$$

On the other hand, the Poisson problem needs a special but a simple consideration due to the point-charges' divergent behavior. The weak form is manipulated as follows,

$$\begin{aligned}
-\frac{1}{4\pi} < \varphi \nabla^2 S_C > &=< \varphi (\rho^{(\mathbf{R}_I)} + b^{(\mathbf{R}_I)}) >, \\
\frac{1}{4\pi} < \nabla \varphi \cdot \nabla S_C > &=< \varphi \rho^{(\mathbf{R}_I)} > + < \nabla \varphi \beta_I^{(0)} >,
\end{aligned} \tag{A.11}$$

by using similarly the integration by parts and gradient theorems. Then, the discretized AE electrostatic problem becomes,

$$\frac{1}{4\pi} \langle \nabla N_I \cdot \nabla N_J \rangle \{S_C\}^J = \langle N_I \rho^{(\mathbf{R}_I)} \rangle + \langle \nabla N_I \beta_I^{(0)} \rangle. \tag{A.12}$$

Validation for AE systems is handled via a DFPT and FD comparison of dihydrogen molecule. The system is modeled with the Perdew-Zunger exchange-correlation functional, LDA. The atom positions are given in Table A.1. The system has a central patch with the dimension of $2.5x2.5x2.5a_0^3$ and an outer patch radius of $50a_0$. Reference energy for the defined system is taken as $E_0 = -1.137844435Ha$, which is calculated with a sixth-order Lagrange discretization corresponding to 435647 DOF.

atomic positions in a_0

atom	x	y	z
H_1	-0.7400	0.0000	0.0000
H_2	0.7400	0.0000	0.0000

Table A.1: Atomic positions for the dihydrogen with LDA in AE setting

The dihydrogen energy converges to $E_0 = -1.137243662Ha$ with a mesh with 72437 DOF. The energy error is in the order of $10^{-4}Ha$. The convergence is not aimed due to the simple mesh construction scheme. Relying on Figure 5.5, DFPT and FD results for IFC are compared directly without complete energy convergence in Figures A.2 and A.3. The IFC results diverge at the order of $10^{-5}Ha$.

$I\alpha$ $J\beta$	11	12	13	21	22	23
11	0.199140	0.000000	0.000000	-0.352157	0.000000	0.000000
12	0.000000	0.447561	0.000000	0.000000	0.014105	0.000000
13	0.000000	0.000000	0.447561	0.000000	0.000000	0.014105
21	-0.352157	0.000000	0.000000	0.199140	0.000000	0.000000
22	0.000000	0.014105	0.000000	0.000000	0.447561	0.000000
23	0.000000	0.000000	0.014105	0.000000	0.000000	0.447561

Table A.2: IFC of dihydrogen with LDA in AE setting by DFPT

$I\alpha$ $J\beta$	11	12	13	21	22	23
11	0.199122	0.000000	0.000000	-0.352147	0.000000	0.000000
12	0.000000	0.447460	0.000000	0.000000	0.014105	0.000000
13	0.000000	0.000000	0.447460	0.000000	0.000000	0.014105
21	-0.352147	0.000000	0.000000	0.199122	0.000000	0.000000
22	0.000000	0.014105	0.000000	0.000000	0.447460	0.000000
23	0.000000	0.000000	0.014105	0.000000	0.000000	0.447460

Table A.3: IFC of dihydrogen with LDA in AE setting by FD

Appendix B

FEM Multiverse

In integro-differential equations, the control over which variable a gradient operates on can simplify both the formulation and computation. It is possible via combinations of the integration by parts with the divergence, gradient, and rotation theorems. However, such manipulations results in surface terms. Fortunately, the surface integrals disappear in variational formulations under homogeneous Dirichlet boundary conditions. With homogeneous Dirichlet boundary conditions, both the trial functions and test functions are constrained to be zero on the boundary. Therefore, it introduces multiple options in the variational FEM.

The last nonlocal contribution term (B.1) in (3.26) is selected for the manipulation. This term naturally derives from (B.2) by considering equivalence of $-\partial/\partial\mathbf{r}$ to $\partial/\partial\mathbf{R}_I$ when acting on $\mathbf{r}_I$.

$$4\sum_{\alpha=1}^{occ/2} f_\alpha \int \nabla\psi_\alpha^{*(0)}(\mathbf{r})\nabla\lambda_I^{(0)}(\mathbf{r})\, d^3\mathbf{r} \int \lambda_I^{*(0)}(\mathbf{r}')\psi_\alpha^{(0)}(\mathbf{r}')\, d^3\mathbf{r}'\delta_{IJ} \tag{B.1}$$

$$4\sum_{\alpha=1}^{occ/2} f_\alpha \int \psi_\alpha^{*(0)}(\mathbf{r})\lambda_{NL}^{(\mathbf{R}_I,\mathbf{R}_J)}(\mathbf{r})\, d^3\mathbf{r} \int \lambda_{NL}^{*(0)}(\mathbf{r}`)\psi_\alpha^{(0)}(\mathbf{r}`)\, d^3\mathbf{r}` \tag{B.2}$$

The multiplications of integrations are domain-independent so that these could be separated for compact inspection. Starting with (B.2), multiverse of formulations are derived by using the integration by parts, gradient theorem, and Dirichlet

boundary conditions. The perturbation indices are transformed to gradients as,

$$\int \psi_{\alpha}^{*(0)} \lambda_{NL}^{(\mathbf{R}_I,\mathbf{R}_J)}\, d^3\mathbf{r} = \int \psi_{\alpha}^{*(0)} \nabla(\nabla\lambda_{NL})\, d^3\mathbf{r}. \tag{B.3}$$

Then, the integration by parts and gradient theorem are enforced as,

$$\int \psi_{\alpha}^{*(0)} \nabla(\nabla\lambda_{NL})\, d^3\mathbf{r} = -\int \nabla\psi_{\alpha}^{*(0)} \nabla\lambda_{NL}\, d^3\mathbf{r} + \int_{\partial\Omega} \psi_{\alpha}^{*(0)} \nabla\lambda_{NL} \otimes \mathbf{n}\, d^2\mathbf{r}, \tag{B.4}$$

where the boundary element vanishes due to the homogeneous Dirichlet boundary conditions. As a result, the same procedure can be expended to the following equivalent formulations,

$$\int \psi_{\alpha}^{*(0)} \nabla(\nabla\lambda_{NL})\, d^3\mathbf{r} = -\int \nabla\psi_{\alpha}^{*(0)} \nabla\lambda_{NL}\, d^3\mathbf{r} = \int \nabla(\nabla\psi_{\alpha}^{*(0)}) \lambda_{NL}\, d^3\mathbf{r}. \tag{B.5}$$

The first and last terms in (B.5) contains second-order gradients in which shape functions complicate the NURBS formulation. The weak form is written as,

$$\int \nabla(\nabla\psi_{\alpha}^{*(0)}) \lambda_I^{(0)}\, d^3\mathbf{r} \approx \int \psi_{\alpha}^{*(0)} \left[\left(\frac{\partial^2 N_I}{\partial\xi^2} \frac{\partial\xi}{\partial\mathbf{r}} \right) \otimes \frac{\partial\xi}{\partial\mathbf{r}} + \left(\frac{\partial N_I}{\partial\xi} \frac{\partial^2\xi}{\partial\mathbf{r}^2} \right) \right] \lambda_I^{(0)}\, d^3\mathbf{r}, \tag{B.6}$$

where

$$\left[\frac{\partial^2 N_I}{\partial\xi_k \partial\xi_l} \frac{\partial\xi_k}{\partial r_i} \frac{\partial\xi_l}{\partial r_j} + \frac{\partial N_I}{\partial\xi_k} \frac{\partial}{\partial r_j} \left(\frac{\partial\xi_k}{\partial r_i} \right) \right] \mathbf{e}_i \otimes \mathbf{e}_j, \tag{B.7}$$

is equivalent to

$$\mathbf{F}^{-T} \frac{\partial^2 N_I}{\partial\xi^2} \mathbf{F}^{-} - \left[\left(\frac{\partial N_I}{\partial\xi} \right)^T \mathbf{F}^{-} \right]^T \frac{\partial\mathbf{F}}{\partial\xi} \mathbf{F}^{-}\mathbf{F}^{-}, \tag{B.8}$$

using the operations,

$$F_{ik}(F^{-})_{kj} = \delta_{ij}, \tag{B.9}$$

$$\begin{aligned} \frac{\partial F_{ik}}{\partial r_j}(F^{-})_{kl} + F_{ik}\frac{\partial (F^{-})_{kl}}{\partial r_j} &= 0 \\ \frac{\partial (F^{-})_{kl}}{\partial r_j} &= -F^{-}_{km}\frac{\partial F_{mn}}{\partial r_j}(F^{-})_{nl}, \end{aligned} \tag{B.10}$$

$$\frac{\partial F_{ik}}{\partial r_j} = \frac{\partial}{\partial r_j}\left(\frac{\partial r_i}{\partial \xi_k}\right) = \frac{\partial^2 r_i}{\partial\xi_k\partial\xi_l}\frac{\partial\xi_l}{\partial r_j}. \tag{B.11}$$

Here, lowercase letters, i.e. $\mathbf{r}$, represent vectors while uppercase letters, i.e. $\mathbf{F}$, represent matrices. At some parts, the formulations are transformed into index form according to the Einstein's notation that double indices mean summation over the index. The unit vectors $\mathbf{e}_i$ is the elements of orthonormal basis set.

The argument detailed with the nonlocal pseudopotential contribution is valid for the atom-atom interactions as well. The atom-atom interaction terms (B.12) in (3.26) is also a result of similar manipulations.

$$-\int\int \frac{\delta\hat{v}_I[\beta_I]}{\delta[\beta_I']}\beta_I^{\prime(\mathbf{R}_I)}\,d^3\mathbf{r}'\beta_I^{(\mathbf{R}_I)}\delta_{IJ}\,d^3\mathbf{r} - \int \hat{v}_I[\beta_I]\beta_I^{(\mathbf{R}_I,\mathbf{R}_I)}\delta_{IJ}\,d^3\mathbf{r} \tag{B.12}$$

The first and the last terms at (B.13) are equivalent. Both converge to the same value as the domain size is selected big enough for the atomic charges to vanish. Either can be preferred for implementation. However, the first term is chosen because the last term requires the second-order perturbation solutions of the atom potential which results in additional Poisson problems. On the other hand, even though the perturbation indices of second and third terms differ, they are equivalent due to the symmetry in atomic charges and potentials. Therefore, both cheaper and neater form given in (B.12) is selected to implement.

$$\begin{aligned}
&-\frac{1}{2}\int \hat{v}_{ext}\beta_I^{(\mathbf{R}_I,\mathbf{R}_I)}\delta_{IJ}\,d^3\mathbf{r} - \frac{1}{2}\int \hat{v}_I^{(\mathbf{R}_I)}[\beta_I]\beta_J^{(\mathbf{R}_J)}\,d^3\mathbf{r} \\
&-\frac{1}{2}\int \hat{v}_J^{(\mathbf{R}_J)}[\beta_J]\beta_I^{(\mathbf{R}_I)}\,d^3\mathbf{r} - \frac{1}{2}\int \hat{v}_I^{(\mathbf{R}_I,\mathbf{R}_I)}[\beta_I]b^{(0)}\delta_{IJ}\,d^3\mathbf{r}.
\end{aligned} \tag{B.13}$$

Appendix C

Finite Difference Scheme

Verification of the IFC tensors via DFPT can be found through second-order derivative on energy and first-order derivative on force. For IFC, positions of nucleus I in the direction of α, $\mathbf{R}_{I\alpha}$, are perturbed in FD scheme. Then, the fourth-order accurate pure FD of second-order derivative scheme, which is applicable to the diagonal elements, is

$$\frac{\partial^2 E}{\partial \mathbf{R}_{I\alpha}^2} \approx \frac{-E_{2\delta\mathbf{R}_{I\alpha}} + 16E_{\delta\mathbf{R}_{I\alpha}} - 30E_0 + 16E_{-\delta\mathbf{R}_{I\alpha}} - E_{-2\delta\mathbf{R}_{I\alpha}}}{12(\delta\mathbf{R}_{I\alpha})^2} + \mathcal{O}((\delta\mathbf{R}_{I\alpha})^4), \tag{C.1}$$

whereas the fourth-order accurate mixed FD of second-order derivative scheme, which is applicable to the off-diagonal elements, is

$$\begin{aligned}\frac{\partial^2 E}{\partial \mathbf{R}_{I\alpha}\mathbf{R}_{J\beta}} \approx & \frac{64\left(E_{\delta\mathbf{R}_{I\alpha},\delta\mathbf{R}_{J\beta}} - E_{\delta\mathbf{R}_{I\alpha},-\delta\mathbf{R}_{J\beta}} - E_{-\delta\mathbf{R}_{I\alpha},\delta\mathbf{R}_{J\beta}} + E_{-\delta\mathbf{R}_{I\alpha},-\delta\mathbf{R}_{J\beta}}\right)}{144(\delta\mathbf{R}_{I\alpha})^2} \\ & + \frac{8\left(E_{2\delta\mathbf{R}_{I\alpha},\delta\mathbf{R}_{J\beta}} - E_{2\delta\mathbf{R}_{I\alpha},-\delta\mathbf{R}_{J\beta}} - E_{-2\delta\mathbf{R}_{I\alpha},\delta\mathbf{R}_{J\beta}} + E_{-2\mathbf{R}_{I\alpha},-\mathbf{R}_{J\beta}}\right)}{144(\delta\mathbf{R}_{I\alpha})^2} \\ & - \frac{\left(E_{3\delta\mathbf{R}_{I\alpha},\delta\mathbf{R}_{J\beta}} - E_{3\delta\mathbf{R}_{I\alpha},-\delta\mathbf{R}_{J\beta}} - E_{-3\delta\mathbf{R}_{I\alpha},\delta\mathbf{R}_{J\beta}} + E_{-3\mathbf{R}_{I\alpha},-\mathbf{R}_{J\beta}}\right)}{144(\delta\mathbf{R}_{I\alpha})^2} \\ & + \mathcal{O}((\delta\mathbf{R}_{I\alpha})^2(\delta\mathbf{R}_{J\beta})^2),\end{aligned} \tag{C.2}$$

where δ imposes the finite FD step on the element it operates.

The fourth-order accurate FD of first-order derivative scheme is

$$\frac{\partial F_{\mathbf{R}_{I\alpha}}}{\partial \mathbf{R}_{J\beta}} \approx \frac{-F_{\mathbf{R}_{I\alpha}+2\delta\mathbf{R}_{J\beta}} + 8F_{\mathbf{R}_{I\alpha}+\delta\mathbf{R}_{J\beta}} - 8F_{\mathbf{R}_{I\alpha}-\delta\mathbf{R}_{J\beta}} + F_{\mathbf{R}_{I\alpha}-2\delta\mathbf{R}_{J\beta}}}{12\delta\mathbf{R}_{J\beta}} + \mathcal{O}((\delta\mathbf{R}_{J\beta})^4), \tag{C.3}$$

where $F_{\mathbf{R}_{I\alpha}}$ is the force component on the ion I in the direction of α.

Appendix D

Vibrational modes

Vibrational modes describe atomic oscillations under a molecular perturbation. These modes are determined by converting IFC (C) to the *dynamical matrix (D)* by mass-weighting the IFC with atomic masses (M) [41].

$$D_{I\alpha,J\beta} = \frac{1}{\sqrt{M_I M_J}} C_{I\alpha,J\beta} \tag{D.1}$$

Each mode is associated with a vibrational frequency (ω) in units of rad/s, which is equivalent to $\sqrt{\frac{Ha}{a_0^2 \cdot amu}}$, and the normalized dimensionless eigenvectors ($\mathbf{e}_{I\alpha}$). The square roots of the dynamical matrix's eigenvalues correspond to the vibrational frequencies while the eigenvectors give atomic displacement patterns, satisfying the eigenvalue equation,

$$D_{I\alpha,J\beta}\mathbf{e}_{I\alpha} = \omega_{I\alpha}^2 \mathbf{e}_{I\alpha}. \tag{D.2}$$

As a convention, the vibrational frequencies in the unit of rad/s are converted to cm^{-1} for comparison with infrared and Raman spectra, where amu denotes the atomic mass unit. Real vibrational frequency indicates stable vibrational modes, which indicates translational and rotational invariance. In contrast, imaginary vibrational frequencies point structural instabilities.

The vibrational frequencies in the unit of cm^{-1} is counterintuitive because it

does not use the unit of frequency (Hz). However, it origins from the Planck-Einstein relation,

$$E = \hbar\omega = 2\pi\hbar c\tilde{v}, \tag{D.3}$$

where $\hbar$ is the reduced Planck constant, c is the speed of light and $\tilde{v}$ corresponds to wavenumber $1/\lambda$ in cm^{-1}. Therefore, the vibrational frequency becomes directly comparable to spectroscopic measurements.

Vibrational modes contain $3N-6$ normal for ordinary molecules while $3N-5$ for linear molecules, where the atoms are co-linear such as H_2, CO_2. Both contain 3 for translational symmetry and the remaining is due to rotational symmetry. A linear molecule of point masses lying in x-axis cannot rotate about its axis so that one of the zero frequencies drops.