Chapter 1

WAVELETS FOR DENSITY-FUNCTIONAL THEORY AND POST-DENSITY-FUNCTIONAL-THEORY CALCULATIONS

Bhaarathi Natarajan[a,b*]
Mark E. Casida[a†]
Luigi Genovese[b‡]
Thierry Deutsch[b§]

[a] Laboratoire de Chimie Théorique,
Département de Chimie Mélécularie (DCM, UMR CNRS/UJF 5250),
Institut de Chimie Moléculaire de Grenoble (ICMG, FR2607),
Université Joseph Fourier (Grenoble I),
301 rue de la Chimie, BP 53,
F-38041 Grenoble Cedex 9, France
[b] UMR-E CEA/UJF-Grenoble 1,
INAC, Grenoble, F-38054, France

October 24, 2011

[*]E-mail address: Bhaarathi.Natarajan@UJF-Grenoble.FR
[†]E-mail address: Mark.Casida@UJF-Grenoble.FR
[‡]E-mail address: Luigi.Genovese@cea.FR
[§]E-mail address: Thierry.Deutsch@CEA.FR

Abstract

We give a fairly comprehensive review of wavelets and of their application to density-functional theory (DFT) and to our recent application of a wavelet-based version of linear-response time-dependent DFT (LR-TD-DFT). Our intended audience is quantum chemists and theoretical solid-state and chemical physicists. Wavelets are a Fourier-transform-like approach which developed primarily in the latter half of the last century and which was rapidly adapted by engineers in the 1990s because of its advantages compared to standard Fourier transform techniques for multiresolution problems with complicated boundary conditions. High performance computing wavelet codes now also exist for DFT applications in quantum chemistry and solid-state physics, notably the BIGDFT code described in this chapter. After briefly describing the basic equations of DFT and LR-TD-DFT, we discuss how they are solved in BIGDFT and present new results on the small test molecule carbon monoxide to show how BIGDFT results compare against those obtained with the quantum chemistry gaussian-type orbital (GTO) based code DEMON2K. In general, the two programs give essentially the same orbital energies, but the wavelet basis of BIGDFT converges to the basis set limit much more rapidly than does the GTO basis set of DEMON2K. Wavelet-based LR-TD-DFT is still in its infancy, but our calculations confirm the feasibility of implementing LR-TD-DFT in a wavelet-based code.

Contents

1.   Introduction

The broad meaning of "adaptivity" is the capacity to make something work better by alternation, modification, or remodeling. Concepts of adaptivity have found widespread use in quantum chemistry, ranging from the construction of Gaussian-type orbital (GTO) basis sets, see e.g., the development of correlation consistent bases [1, 2, 3], to linear scaling methods in density functional theory (DFT) [4, 5, 6, 7, 8, 9], selective configuration interaction (CI) methods [10, 11] and local correlation methods based on many-body perturbation theory or coupled cluster (CC) theory [12, 13]. This chapter is about a specific adaptive tool, namely wavelets as an adaptive basis set for DFT calculations which can be automatically placed when and where needed to handle multiresolution problems with difficult boundary conditions.

Let us take a moment to contrast the wavelet concept of adaptivity with other types of adaptivity. In other contexts, the adaptive procedure is typically based on a combination of physical insights together with empirical evidence from numerical simulations. A rigorous mathematical justification is usually missing. This may not be surprising: Familiar concepts lose a lot of their original power if one tries to put them in a rigorous mathematical framework. Therefore, we will not shoulder the monumental and perhaps questionable task of providing a rigorous mathematical analysis of all the adaptive approaches used nowadays throughout quantum chemistry. Instead we will concentrate on the mathematical analysis of a particular electronic structure method which lends itself to a rigorous mathematical analysis and application of adaptivity. In contrast with other adaptive methods, multiresolution analysis (MRA) with wavelets can be regarded as an additive subspace correction and their wavelet representations have a naturally built-in adaptivity which comes through

their ability to express directly and separate components of the desirable functions living on different scales.

This combined with the fact that many operators and their inverses have nearly sparse representations in wavelet coordinates may eventually lead to very efficient schemes that rely on the following principle: Keep the computational work proportional to the number of significant coefficients in the wavelet expansions of the searched solution. As there are a lot of different wavelet bases with different properties (length of support, number of vanishing moments, symmetry, etc.) in each concrete case we can choose the basis that is most appropriate for the intrinsic complexity of the sought-after solution. This fact makes the wavelet-based schemes a very sophisticated and powerful tool for compact representations of rather complicated functions. The expected success of wavelet transforms for solving electronic structure problems in quantum mechanics are due to three important properties: (a) the ability to choose a basis set providing good resolution where it is needed, in those cases where the potential energy varies rapidly in some regions of space, and less in others; (b) economical matrix calculations due to their sparse and banded nature; and (c) the ability to use orthonormal wavelets, thus simplifying the eigenvalue problem. Of course, this might lead to adaptive methods which are fully competitive from a practical point of view, for example, working with a systematic basis instead of GTO bases requires from the onset larger basis sets and the benefit of systematic improvement might be a distant prospect. However, we have the more realistic prospect that our rigorous analysis provides new and hopefully enlightening perspectives on standard adaptive methods, which we reckon cannot be obtained in another way.

On the otherhand advances in computational technology opened up new opportunities in quantum mechanical calculation of various electronic structures, like molecules, crystals, surfaces, mesoscopic systems, etc. The calculations can only be carried out either for very limited systems or with restricted models, because of their great demand of computational and data storage resources. Independent particle approximations, like the Hartree-Fock based [14, 15, 16, 17] algorithms with single determinant wave functions, leave out the electron correlation and need operation and storage capacity of order N^4, if N is the total number of electrons in the system. If inclusion of the electron correlation is necessary, CI or CC methods can be applied, with very high demand of computational resources ($\mathcal{O}(N^6)$ to $\mathcal{O}(N!)$). An alternative way is to use MBPT. The second order perturbation calculations can be carried out within quite reasonable time and resource limits, but the results are usually unsatisfactory, they just show the tendencies, while the 4th order MBPT needs $\mathcal{O}(N^7)$ to $\mathcal{O}(N^8)$ operations. All these algorithms use the N-electron wave function as a basic quantity.

Another branch of methods use electron density as the primary entity. Pioneers of this trend, like Thomas [18], Fermi [19, 20], Frenkel [21] and Sommerfeld [22] developed the statistical theory of atoms and the local density approximation (LDA). The space around the nuclei is separated into small regions, where the atomic potential is approximated as a constant, and the electrons are modeled as a free electron gas of Fermi-Dirac statistics [23, 24, 21]. Dirac included electron correlation [25], which improved the results. After the Hohenberg–Kohn theorems had appeared [26], and Kohn and Sham had offered a practically applicable method [27] based on their work, many scientists were motivated to work on the theory, and DFT developed into one of the most powerful electronic structure

methods.

Despite the success of density functional theory, it has some drawbacks. The exact formula of the exchange-correlation potential is not known, thus chemical intuition and measured data are necessary in order to approximate it, and the kinetic energy functional is hard to calculate. Powerful approximating formulas are available (see, e.g. [28]), like the Thomas–Fermi functional based gradient and generalized gradient expansions, where the energy functionals are expressed as a power series of the gradient of the density (the first such suggestion was [29].)

Considering the historical development of sophisticated N-electron methods, a typical trend can be observed. Starting with a very simple model, new details are introduced in order to improve the results. This scheme is followed in the linear muffin tin orbital method (LMTO) [30] where the interatomic regions is replaced by the spherical orbital of an atomic potential around the nuclei. Similarly, the linearized augmented plane wave method (APW) [31] and the plane wave pseudopotential approach [32] describe the details of the crystal potential differently in different spatial regions. Although they are rather successful, for applying any of these models, chemical intuition is needed, free parameters, like the radius of the bordering sphere between the two types of potentials, and the boundary conditions have to be set. A systematic method, which can handle the different behaviors of the electron structures at different spatial domains, or either at different length scale [33], is the longterm requirement of any physical chemists.

Multiresolution or wavelet analysis, this rapidly developing branch of the applied mathematics, is exactly the tool for statisfy all the need of any chemical physicists/physcial chemists. From mathematical point of view wavelet analysis is a theory of a special kind of Hilbert space basis sets. Basis sets are commonly used in all electron structure calculations, as the wave function is usually expanded as linear combination of some kind of basis functions. Thus the operator eigenvalue problem is reduced to an algebraic matrix eigenvector problem. The resulting algebraic equations are easier to solve, well known algorithms and subroutine libraries are available, however, the difficulty of choosing the proper basis set arises. If linear combination of atomic orbitals (LCAO) is used, the atomic basis functions are Slater or Gaussian-type of functions [34, 35], the selection of atomic orbitals needs chemical intuition, which is a result of long time's experience, and can not be algorithmized. Both basis sets are non-orthogonal, and lack the explicit convergence properties [36]. Moreover, calculation of operator matrix elements with Slater-type orbitals is complicated, their integrals have to be treated numerically. Although integrals of Gaussian functions are analytically known, the Gaussisn-type basis does not reflect the nuclear cusp condition of Kato [37], which reflects on singularities of the N-electron wave function in the presence of Coulomb-like potentials. Since then it turned out that for high precision numerical calculations it is essential to satisfy these requirements. However, while the nuclear cusp condition is relatively easy to fulfill by Slater-type orbitals (STO), the electron-electron cusp is extremely hard to represent. In general, GTO-based/STO-based DFT codes gives reliable results with a relatively small number of basis functions, making them optimal for large scale computations where high accuracy is less crucial. On the other hand there is no consistent way to extend these basis sets and thereby converge the results with respect to the size of the basis.

The second type of basis set covers the system-independent functions such as plane

waves [32] or wavelets [38]. The main advantable of these basis sets, is that their size can be systematically increased until the result of the calculation has converged, and are generally considered to be more accurate than the former type. The number of basis function required to obtain convergence is normally *so large* that direct solution of the matrix eigenvalue problem within the entire basis space is *not* possible. Instead one has to use iterative methods to determine the lowest (occpied) part of the spectrum [32]. In solid state physics, where more or less periodic systems are studied, choosing plane wave basis sets is rather usual. These basis functions are system independent and easily computable, but the results are not always convincing and the number of necessary basis functions is almost untreatable. (Theoretically, plane waves could also be used for describing molecules, since the two-electron integrals and the expectation values are connected to the Fourier transform, thus they are easily computable, and this could balance the large number of necessary basis functions.) The reason, why so many plane waves are needed is that the wave functions around the nuclei need very high frequency terms, i.e. high resolution level, for reproducing the nuclear cusps. In the framework of Fourier analysis, the whole space has to be expanded at the same resolution, despite that in most of the space low frequency terms would be sufficient.

Fully-numerical "basis-set free" Hartree-Fock (HF) calculations of atoms have been known since the 1960s (Vol. 1, pp. 322-326 and Vol. 2, pp. 15-30 of Ref. [39] and Refs. [40, 41, 42, 43]) and have proven helpful in constructing efficient finite basis sets for molecular calculations. In the late 1980s, Axel Becke used a fully-numerical density-functional theory (DFT) program for diatomics to show that many of the problems of DFT calculations at that time were due not to the functionals used, but rather numerical artifacts of the DFT programs of the 1970s [44].) Since that time, fully-numerical DFT codes have been implemented for polyatomic molecules using the finite element method (FEM), with PARSEC from the chemists point of view or OCTOPUS from the view of physicsts being a notable example.

BIGDFT the pseudo potential code for bigger systems based as it is on traditional Hohenberg-Kohn-Sham DFT [26, 27], could only calculate ground-state properties with an eye to order-N DFT. As a step to increase the feasibility of the code we formulated the wavelet-based linear-response time-dependent density-functional theory (TD-DFT) and here we support our first implementation for calculating electronic excitation spectra [45]. Electronic excitation spectra can be calculated from TD-DFT [46] using time-dependent linear response (LR) theory [47, 48]. Casida formulated LR-TD-DFT (often just refered to as TD-DFT) so as to resemble the linear-response time-dependent HF equations already familiar to quantum chemists [48]. That method was then rapidly implemented in a large number of electronic structure codes in quantum chemistry, beginning with the DEMON family of programs [49] and the TURBOMOL program [50]. Among the programs that implemented "Casida's equations" early on was the FEM DFT program PARSEC [51] and also be found in the FEM DFT program OCTOPUS [52]. See Ref. [53] for a recent FEM implementation of TD-DFT. Since a wavelet-based program offers certain advantages over these other FEM DFT programs, it was deemed important to also implement LR-TD-DFT in BIGDFT.

In the next section we give a detailed description of the idea behind the multiresolution analysis and wavelets, with a historical note. Sec. 3. and Sec. 4., briefly presenting the the-

oretical introduction to DFT and TD-DFT, and Sec. 5., talks about the well-known Krylov space methods for solving eigenvalue equations involved in our implementation. Sec. 6. and Sec. 7., gives the numerical implementation of DFT and how we have implemented TD-DFT from the aspects of theoretical and algorithmic point of view on wavelets based pseudopotential electronic structure code BIGDFT, and in Section 8. we give the results of detailed comparisons between TD-DFT excitation spectra calculated with BIGDFT and with the implementation of Casida's equations in the GTO-based program DEMON2K. The conclusion were drawn for future applications in the field of chemistry and some of the other problems are reviewed to draw chemists' greater attention to wavelets and to gain more benifits from using wavelet technique.

2.   Wavelet Theory

The mathematics of wavelets is a fairly new technique, it can generally be used where one traditionally uses Fourier techniques. They incorporate the feature of having multiple scales, so very different resolutions can be used in different parts of space in a mathematically rigorous manner. This matches many systems in nature well, for example molecule where the atomic orbitals are very detailed close to the cores, while they only vary slowly between them. Wavelet analysis can quite generally be viewed as a local Fourier analysis. From the wavelet expansion, or wavelet spectrum, of a function, f, it can be inferred not only how fast f varies, i.e. which frequencies it contains, but also where in space a given frequency is located. This property has important applications in both data compression, signal/image processing and noise reduction [54]. Wavelet methods are also employed for solving partial differential equations [55, 56], and in relation to electronic structure methods a complete DFT program based on interpolating wavelets has been developed [57].

2.1.   The story of wavelets

Most historical versions of wavelet theory however, despite their source's perspective, begin with Joseph Fourier. In 1807, a French mathematician, Joseph Fourier, discovered that all periodic functions could be expressed as a weighted sum of basic trigonometric functions. His ideas faced much criticism from Lagrange, Legendre and Laplace for lack of mathematical rigor and generality, and his papers were denied publication. It took Fourier over 15 years to convince them and publish his results. Over the next 150 years his ideas were expanded and generalized for non-periodic functions and discrete time sequences. The fast Fourier transform algorithm, devised by Cooley and Tukey in 1965 placed the crown on Fourier transform, making it the king of all transforms. Since then Fourier transforms have been the most widely used, and often misused, mathematical tool in not only electrical engineering, but in many disciplines requiring function analysis. This crown however, was about to change hands. Following a remarkably similar history of development, the wavelet transform is rapidly gaining popularity and recognition.

The first mention of wavelets was in a 1909 dissertation by Hungarian mathematician Alfred Haar. Haar's work was not necessarily about wavelets, as "wavelets" would not appear in their current form until the late 1980s. Specifically, Haar focused on orthogonal function systems, and proposed an orthogonal basis, now known as the Haar wavelet basis,

in which functions were to be transformed by two basis functions. One basis function is constant on a fixed interval, and is known as the scaling function. The other basis function is a step function that contains exactly one zero–crossing (vanishing moment) over a fixed interval (more on this later).

The next major contribution to wavelet theory was from a 1930s French scientist Paul Pierre Lévy. More correctly, Lévy's contribution was less of a contribution and more of a validation. While studying the ins and outs of Brownian motion in the years following Haar's publication, Lévy discovered that a scale–varied Haar basis produced a more accurate representation of Brownian motion than did the Fourier basis. Lévy, being more of a physicist than mathematician, moved on to make large contributions to our understanding of stochastic processes.

Contributions to wavelet theory between the 1930s and 1970s were slight. Most importantly, the windowed Fourier transform was developed, with the largest contribution being made by another Hungarian named Dennis Gabor. The next major advancement in wavelet theory is considered to be that of Jean Morlet in the late 1970s.

Morlet, a French geophysicist working with windowed Fourier transforms, discovered that fixing frequency and stretching or compressing (scaling) the time window was a more useful approach than varying frequency and fixing scale. Furthermore, these windows were all generated by dilation or compression of a prototype Gaussian. These window functions had compact support both in time and in frequency (since the Fourier transform of a Gaussian is also a Gaussian.) Due to the small and oscillatory nature of these window functions, Morlet named his functions as "wavelets of constant shape". In 1981, Morlet worked with Croatian–French physicist Alex Grossman on the idea that a function could be transformed by a wavelet basis and transformed back without loss of information, thereby outlining the wavelet transformation. It is of note that Morlet initially developed his ideas with nothing more than a handheld calculator.

In 1986, Stéphane Mallat noticed a publication by Yves Meyer that built on the concepts of Morlet and Grossman. Mallat sought Meyer's consult, and the result of said consult was Mallat's publication of multiresolution analysis. Mallat's MRA connected wavelet transformations with the field of digital signal processing. Specifically, Mallat developed the wavelet transformation as a multiresolution approximation produced by a pair of digital filters. The scaling and wavelet functions that constitute a wavelet basis are represented by a pair of finite impulse response filters, and the wavelet transformation is computed as the convolution of these filters with the input function. The importance of Mallat's contribution cannot be overstated. Without the fast computational means of wavelet transformation provided by the MRA, wavelets, undoubtedly, would not be the effective and widely used signal processing tools that they are today.

In 1988, a student of Alex Grossman, named Ingrid Daubechies, combined the ideas of Morlet, Grossman, Mallat, and Meyer by developing the first family of wavelets as they are known today. Named the Daubechies wavelets, the family consists of 8 separate wavelet and scaling functions (more on this later). With the development of pair Daubechies wavelet and scaling functions is orthogonal, continuous, regular, and compactly supported, the foundations of the modern wavelet theory were laid. The last ten years mostly witnessed a search for other wavelets with different properties and modifications of the MRA algorithms. In 1992, Albert Cohen, Jean Feauveau and Daubechies constructed the compactly supported

biorthogonal wavelets, which are preferred by many researchers over the orthonormal basis functions, whereas R. Coifman, Meyer and Victor Wickerhauser developed wavelet packers, a natural extension of MRA.

2.2. Multiresolution analysis

A suitable gateway to the theory of wavelets is through the idea of MRA. A detailed description of MRAs can be found in Keinert [58], from which a brief summary of the key issues are given in the following.

A multiresolution analysis is an infinite nested sequence of subspaces $L^2(\mathbb{R})$

$$V_j^0 \subset V_j^1 \subset ... \subset V_j^n \subset ... \tag{1}$$

with the following properties

- V_j^∞ is dense in L^2

- $f(x) \in V_j^n \iff f(2x) \in V_j^{n+1} 0 \leq n \leq \infty$

- $f(x) \in V_j^n \iff f(x - 2^{-n}l) \in V_j^n 0 \leq l \leq (2^n - 1)$

- There exists a function vector φ of length $j + 1$ in L^2 such that

$$\{\varphi_j(x) : 0 \leq k \leq j\}$$

forms a basis for V_j^0.

This means that if we can construct a basis of V_j^0, which consists of only $j+1$ functions, we can construct a basis of any space V_j^n, by simple compression (by a factor of 2^n), and translations (to all grid points at scale n), of the original $j + 1$ functions, and by increasing the scale n, we are approaching a complete basis of L^2. Since $V_j^n \subset V_j^{n+1}$ the basis functions of V_j^n can be expanded in the basis of V_j^{n+1}

$$\varphi_l^n(x) \stackrel{\text{def}}{=} 2^{n/2}\varphi(2^n x - l) = \sum_l h^{(l)} \varphi_l^{n+1}(x) \,. \tag{2}$$

where $h^{(l)}$s are the so-called filter matrix that describes the transformation between different spaces V_j^n.

The MRA is called orthogonal if

$$\langle \varphi_0^n(x), \varphi_l^n(x) \rangle = \delta_{0l} I_{j+1} \,, \tag{3}$$

where I_{j+1} is the $(j + 1) \times (j + 1)$ unit matrix, and $j + 1$ is the length of the function vector. The orthogonality condition means that the functions are orthogonal both within one function vector and through all possible translations on one scale, but not through the different scales.

Complementary to the nested sequence of subspaces V_j^n, we can define another series of spaces W_j^n that complements V_j^n in V_j^{n+1}

$$V_j^{n+1} = V_j^n \oplus W_j^n \tag{4}$$

where there exists another function vector ϕ of length $j + 1$ that, with all its translations on scale n forms a basis for W_j^n. Analogously to Eq. (2) the function vector can be expanded in the basis of V_j^{n+1}

$$\phi_l^n(x) \stackrel{\text{def}}{=} 2^{n/2}\phi(2^n x - l) = \sum_l g^{(l)}\phi_l^{n+1}(x).$$ (5)

with filter matrices $g^{(l)}$. The functions ϕ also fulfill the same orthogonality condition as Eq. (3), and if we combine Eq. (1) and Eq. (4) we see that they must be orthogonal with respect to different scales. Using Eq. (4) recursively we obtain

$$V_j^n = V_j^0 \oplus W_j^0 \oplus W_j^1 \oplus ... \oplus W_j^{n-1}.$$ (6)

which will prove to be an important relation.

2.3. Wavelets

There are many ways to choose the basis functions φ and ϕ (which define the spanned spaces V_j^n and W_j^n), and there have been constructed functions with a variety of properties, and we should choose the wavelet family that best suits the needs of the problem we are trying to solve. (Wavelets are often denoted by ψ in the literature but the choice has been made here to denote them by ϕ so as to avoid confusion with the Kohn-Sham orbitals.) Otherwise, we could start from scratch and construct the new family, one that is custum-made for the problem at hand. Of course, this is not a trivial task, and it might prove more efficient to use an existing family, even though its properties are not right on cue.

There is a one-to-one correspondence between the basis functions φ and ϕ, and the filter matrices $h^{(l)}$ and $g^{(l)}$ used in the two-scale relation equations Eq. (2) and Eq. (5), and most well-known wavelet families are defined only through their filter coefficients.

In the following we are taking a different, more intuitive approach, for defining the *scaling space* V_j^n as the space of piecewise polynomial functions

$$V_j^n \quad \stackrel{\text{def}}{=} \quad \{f : \text{ all polynomials of degree} \leq j$$

$$\text{on the interval}(2^{-n}l, 2^{-n}(l+1))$$
$$\text{for} 0 \leq l < 2^n, f \text{ vanishes elsewhere}\}.$$ (7)

It is quite obvious that one polynomial of degree j on the interval $[0, 1]$ can be exactly reproduced by two polynomials of degree j, one on the interval $[0, \frac{1}{2}]$ and the other on the interval $[\frac{1}{2}, 1]$. The spaces V_j^n hence fulfills the MRA condition Eq. (1), and if the polynomial basis is chosen to be orthogonal, the V_j^n constitutes an orthogonal basis.

2.4. An example: Simple Haar wavelets

The basic wavelet ideas that we need can be easily explained using Haar wavelets [59]. These are simply the box functions shown in Fig. 2. We begin with a compact "mother
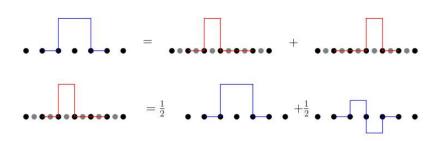
Figure 1. Wavelets (bottom) and scaling function (top).

scaling function," in this case the Haar function,

$$\varphi(x) = \left\{ \begin{array}{lll} 0 & ; & x > 1 \\ 1 & ; & 0 < x < 1 \\ 0 & ; & x < 0 \end{array} \right. . \tag{8}$$

Translations, $\{\varphi_i(x) = \varphi(x - i)\}$, of this mother function produces a crude basis set. Its relation to the grid of integers is obvious. Successively more refined basis sets may be generated by repeated application of the scaling operation consisting of contracting the functions to half their size in the x direction. The kth generation of scaling function is given by $\left\{ \varphi_i^{(k)}(x) = \varphi(2^k x - i) \right\}$. Each generation has a fixed resolution related to an underlying grid with the same resolution. Let us now try to construct a multiresolution basis set. This is accomplished by (say) beginning with the third generation wavelets and taking sums and differences of adjacent functions until the eight third generation scaling functions have been replaced with the eight wavelets shown in Fig. 3. Notice how each generation of daughter wavelets is related to the mother wavelet by scaling, $\phi_i^{(k)}(x) = \phi(2^k x - i)$. Notice also how the mother and two generations of daughter wavelets plus the mother scaling function (occasionally refered to as the "father wavelet") constitute a multiresolution basis set equivalent to the original third generation scaling basis set. Thus an arbitrary function, $f(x)$, expressible in the original scaling basis,

$$f(x) = \sum_{i=1}^{8} \varphi_i^{(3)}(x) s_i^{(3)} , \tag{9}$$

has the wavelet transform,

$$f(x) = \varphi_0^{(0)}(x) s_0^{(0)} + \phi_0^{(0)}(x) d_0^{(0)} + \sum_{i=0,1} \phi_i^{(1)}(x) d_i^{(1)} + \sum_{i=0,3} \phi_i^{(2)}(x) d_i^{(2)} . \tag{10}$$

Since the basis set is multiresolution, we may choose to add more grid points in some region of space and go locally to higher order wavelet expansions. It is also not always necessary to carry out a full wavelet transform, but rather it may be useful to just carry out a partial transform giving a linear combination of wavelets with several scaling functions at a time.
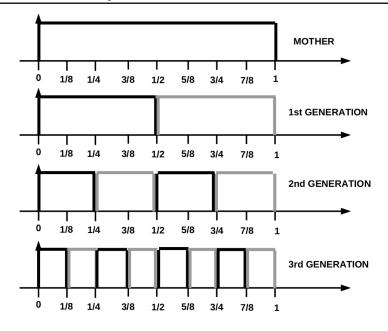
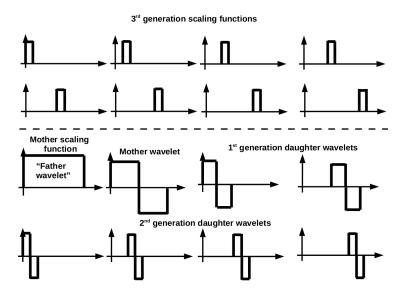Figure 2. Haar scaling functions.

Figure 3. Haar scaling functions and the corresponding wavelets.

The extension to three dimensions is accomplished by using products of one-dimensional scaling functions and wavelets. Haar wavelets are just one type of wavelet basis set. It happens to be pedagogically useful but is not particularly useful for computations.

2.5. Wavelet Basis

The *wavelet space* W_j^n is defined, according to Eq. (4), as the orthogonal complement of V_j^n in V_j^{n+1}. The wavelet basis functions of W_j^n are hence piece-wise polynomials of degree $\leq j$ on each of the two intervals on scale $n+1$ that overlaps with one interval on scale n. These piece-wise polynomials are then made orthogonal to a basis of V_j^n and to each other. The construction of the wavelet basis follows exactly [60] where a simple Gram-Schmidt orthogonalization were employed to construct a basis that met the necessary orthogonality conditions.

One important property of the wavelet basis is the number of vanishing moments. The j-th continuous moment of a function ϕ is defined as the integrals

$$\mu_j \stackrel{\text{def}}{=} \int_0^1 x^j \phi(x) dx \,, \tag{11}$$

and the function ϕ has M vanishing moments if

$$\mu_j = 0, \quad k = 0, ..., M-1$$

The vanishing momenets of the *wavelet* functions gives information on the approximation order of the *scaling* functions. If the wavelet function ϕ has M vanishing moments, any polynomial of order $\leq M-1$ can be exactly reproduced by the scaling function φ, and the error in representing an arbitrary function in the scaling basis is of M-th order. By construction, x^i is in the space V_j^0 for $0 \leq i \leq j$, and since $W_j^0 \perp V_j^0$, the first $k+1$ moments of ϕ_j^0 must vanish.

2.6. The scaling basis

The construction of the scaling functions is quite straightforward, $j+1$ suitable polynomials are chosen to span any polynomial of degree j on the unit interval. The total basis for V_j^n is then obtained by appropriate dilation and translation of these functions. Of course, any polynomial basis can be used, the simplest of them the standard basis $\{1, x, ..., x^j\}$. However, this basis is not orthogonal on the unit interval. In the following, two choices of orthogonal scaling functions will be presented, and even though they span exactly the same spaces V_j^n there are some important numerical differences between the two.

In order to construct a set of orthogonal polynomials we could proceed in the same manner as for the wavelet functions and do a Gram-Schmidt orthogonalization of the standard basis $\{1, x, ...x^j\}$. If this is done on the interval $x \in [-1, 1]$ we end up with the Legendre polynomials $\{L_k\}_{k=0}^j$. These functions are usually normalized such that $L_k(1) = 1$ for all j. To make the *Legendre scaling functions* φ_k^L we transform the Legendre polynomials to the interval $x \in [0, 1]$, and L^2 normalize

$$\varphi_k^L(x) = \sqrt{2k+1} L_k(2x - 1), \quad x \in [0, 1] \,. \tag{12}$$

The basis for the space V_j^n is then made by proper dilation and translation of φ_k^L. Alpert et al. [60] presented an alternative set of scaling functions with interpolating properties. These *interpolating scaling functions* φ_k^I are based on the Legendre scaling functions $\{\varphi_k^L\}_{k=0}^j$, and the roots $\{y_k\}_{k=0}^j$ and weights $\{w_k\}_{k=0}^j$ of the Gauss-Legendre quadrature of order $j+1$, and are organized in the linear combinations

$$\varphi_k^I(x) = \sqrt{w_k} \sum_{i=0}^{j_p} \varphi_i^L(y_k)\varphi_i^L(x), \quad x \in [0,1], \tag{13}$$

Again the basis of V_j^n is made by dilation and translation of ϕ_k^I. The construction of φ_k^I gives them the interpolating property

$$\varphi_k^I(y_i) = \frac{\delta_{ki}}{\sqrt{w_i}} . \tag{14}$$

which will prove important for numerical efficiency.

A detailed discussion on the properties of interpolating wavelets can be found in Donoho [61], but the case of interpolating wavelets is somewhat different. An important property of interpolating wavelets is the smoothness of any function represented in this basis. This property stems from general Lagrange interpolation. In the wavelet case the interpolating property applies within one scaling function vector only, which means that functions represented in this basis can be discontinous in any merging point between the different translations on any scale.

2.7. Interpolating scaling functions

Since the general introduction to wavelets were already made, we will now concencentrate our description on the level 3 interpolating scaling function (ISF) introduced by Deslauriers and Dubuc, and described in detail in Ref. [62]. Its main advantage is that it is fast and easy to perform nonlinear operations on functions represented in this basis, as long as the operation is local in shape. It also represents 3rd order polynomials exactly which means that it behaves very smoothly.

We introduced the projection operator P^n that projects an arbitrary function $f(x)$ onto the basis $\{\varphi_{j,l}^n\}$ of the scaling space V^n (in the remaining of this text the subscript k of the scaling and wavelet spaces will be omitted, and it will always be assumed that we are dealing with a kth order polynomial basis.)

$$f(x) \approx P^n f(x) \stackrel{\text{def}}{=} f^n(x) = \sum_{l=0}^{2^n-1} \sum_{j=0}^{k} s_{j,l}^{n,f} \varphi_{j,l}^n(x) , \tag{15}$$

where the expansion coefficients $s_{j,l}^{n,f}$, the so-called *scaling* coefficients, are obtained by the usual integral

$$s_{j,l}^{n,f} \stackrel{\text{def}}{=} \langle f, \varphi_{j,l}^n \rangle = \int_0^1 f(x)\varphi_{j,l}^n(x)dx , \tag{16}$$

If this approximation turns out to be too crude, we double our basis set by increasing the scale and perform the projection P^{n+1}. This can be continued until we reach a scale N where we are satisfied with the overall accuracy of f^N relative to the true function f.

In a perfect world, the projection in Eq. (16) could be done exactly, and the accuracy of the projection would be independent of the choice of polynomial basis. In the real world the projections are done with Gauss-Legendre quadrature and the expansion coefficients $s_{j,l}^{n,f}$ of $f(x)$ are obtained as

$$
\begin{aligned}
s_{j,l}^{n,f} &= \int_{2^{-nl}}^{2^{-n(l+1)}} f(x)\varphi_{j,l}^{n}(x)dx \\
&= 2^{-n/2} \int_{0}^{1} f(2^{-n}(x+l))\varphi_{j,0}^{0}(x)dx \\
&\approx 2^{-n/2} \sum_{q=0}^{k_q-1} w_q f(2^{-n}(y_q+l))\varphi_{j,0}^{0}(y_q)
\end{aligned}
\tag{17}
$$

where $\{w_q\}_{q=0}^{k_q-1}$ are the weights and $\{y_q\}_{q=0}^{k_q-1}$ the roots of the Legendre polynomial L_{k_q} used in k_qth order quadrature.

By approximating this integral by quadrature we will of course not obtain the exact expansion coefficients. However, it would be nice if we could obtain the exact coefficients whenever our basis is flexible enough to reproduce the function exactly, that is if $f(x)$ is a polynomial of degree $\leq k$. The Legendre quadrature holds a $(2k-1)$ rule which states that the k-order quadrature is exact whenever the integrand is a polynomial of order $2k-1$. By choosing $k_q = k+1$ order quadrature we will obtain the exact coefficient whenever $f(x)$ is a polynomail of degree $\leq (k+1)$ when projecting on the basis of k-order Legendre polynomials.

In the multidimensional case the expansion coefficients are given by multidimensional quadrature

$$
s_{jl}^{nf} = 2^{-dn/2} \sum_{q_1=0}^{k} \sum_{q_2=0}^{k} ... \sum_{q_d=0}^{k} f(2^{-n}(y_q+l))\Pi_{i=1}^{d} w_{q_i}\varphi_{j_p,0}^{0}(y_{q_i}),
\tag{18}
$$

using the following notation for the vector of quadrature roots

$$
y_q \stackrel{\text{def}}{=} (y_{q_1}, y_{q_2}, ..., y_{q_d}),
\tag{19}
$$

This quadrature is not very efficient in multiple dimensions since the number of terms scales as $(k+1)^d$. However, if the function f is separable and can be written $f(x_1, x_2, ..., x_d) = f_1(x_1)f_2(x_2)...f_d(x_d)$, Eq. (18) can be simplified to

$$
s_{jl}^{nf} = 2^{-dn/2}\Pi_{i=1}^{d} \sum_{q_i=0}^{k} f_i(2^{-n}(y_{q_i}+l_i))w_{q_i}\varphi_{j_i,0}^{0}(y_{q_i}),
\tag{20}
$$

which is a product of small summations and scales only as $d(k+1)$.

The Legendre polynomials show very good convergence for polynomial functions $f(x)$, and are likely to give more accurate projections. However, most interesting functions $f(x)$ are not simple polynomials, and the accuracy of the Legendre scaling functions versus a general polynomial basis might not be very different.

By choosing the quadrature order to be $k+1$ a very important property of the Interpolating scaling functions emerges, stemming from the specific construction of these functions Eq. (13), and the use of the $k + 1$ order quadrature roots and weights. The interpolating property Eq. (14) inserts a Kronecker delta whenever the scaling function is evaluated in a quadrature root, which is exactly the case in the quadrature sum. This reduces Eq. (17) to

$$s_{jl}^{n,f} = \frac{2^{-n/2}}{\sqrt{w_j}} f(2^{-n}(x_j + l)), \tag{21}$$

which obviously makes the projection $k + 1$ times more efficient.

In multiple dimensions this property becomes even more important, since it effectively removes all the nested summations in Eq. (18) and leaves only one term in the projection

$$s_{jl}^{nf} = f(2^{-n}(y_j + l))\Pi_{i=1}^{d}\frac{2^{-n/2}}{\sqrt{w_{j_i}}}, \tag{22}$$

This means that in the Interpolating basis the projection is equally effective regardless of the separability of the function f.

3. Density Functional Theory

A method to resolve the electronic structure is by using variational principle

$$E[\Psi] = \frac{\langle\Psi|\hat{H}|\Psi\rangle}{\langle\Psi|\Psi\rangle}, \tag{23}$$

Where $\langle\Psi|\hat{H}|\Psi\rangle = \int d\mathbf{r}\Psi^*(\mathbf{r})\hat{H}\Psi(\mathbf{r})$, Ψ denotes the electronic wavefunction and $\hat{H}$ the Hamiltonian. The energy computed from a guess Ψ is an upper bound to the true ground state energy E_0. Full minimization of the functional $E[\Psi]$ will give the true ground state Ψ^{gs} and energy $E_0 = E[\Psi^{gs}]$.

Density-functional theory states that the many electron problem can be replaced by an equivalent set of self-consistent one-electron equations, the Kohn-Sham equations

$$\hat{h}\psi_i^\sigma(\mathbf{r}) = \left(-\frac{1}{2}\nabla^2 + \hat{v}_{pp}(\mathbf{r}) + \hat{v}_H(\mathbf{r}) + \hat{v}_{xc}^\sigma(\mathbf{r})\right)\psi_i^\sigma(\mathbf{r}) = \epsilon_i^\sigma\psi_i^\sigma(\mathbf{r}). \tag{24}$$

The eigenfunctions ψ_i^σ are the one-electron wavefunctions that correspond to the minimum of the Kohn-Sham energy functional. In these wavefunctions, i is the orbital index and σ denotes the spin, which can be either up $\uparrow$ or down $\downarrow$ (spin α or β.)

The Hamiltonian $\hat{H}$ consists of four different parts: a part related to the kinetic energy of the electrons, the pseudopotentials $\hat{v}_{psp}$, the Hartree potential $\hat{v}_H$ and the exchange correlation potential $\hat{v}_{xc}$. The interaction of the positively charged nuclei with the electrons is described using the pseudopotential $\hat{v}_{psp}$ instead of using the full Coulombic potential. The pseudopotential usually consists of both a local and a non-local part

$$\hat{v}_{psp}(\mathbf{r}) = v_{loc}(r) + \sum_l |l\rangle\hat{v}_l(r,r')\langle l|. \tag{25}$$

The Hartree potential $\hat{v}_H$ describes the interaction between electrons and is given by

$$\hat{v}_H(\mathbf{r}) = \int d\mathbf{r}' \frac{\rho_\uparrow(\mathbf{r}) + \rho_\downarrow(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \,. \tag{26}$$

Finally, the exchange correlation potential $\hat{v}_{xc}$ describes the nonclassical interaction between the electrons and is given by the functional derivative of an exchange correlation energy functional

$$\hat{v}_{xc}^\sigma(\mathbf{r}) = \frac{\delta E_{xc}(\rho_\uparrow, \rho_\downarrow)}{\delta \rho_\sigma(\mathbf{r})} \,. \tag{27}$$

In these equations ρ^σ is the electron spin density, defined as

$$\rho_\sigma(\mathbf{r}) = \sum_i n_i^\sigma |\psi_i^\sigma(\mathbf{r})|^2 \,, \tag{28}$$

where n_i^σ is the occupation number, i.e. the number of electrons in orbital i. In case of LDA (which we use throughtout this chapter) where there is no longer a distinction between spin up and spin down, orbitals can contain at most two electrons.

4. Time-Dependent Density Functional Theory

This section contains a brief review of the basic formalism of TD-DFT which is already well-known from the literature [46]. The time-dependent single particle Kohn-Sham equations are,

$$\left(-\frac{1}{2} \nabla^2 + v_{eff}[\rho](\mathbf{r}, t) \right) \psi_{i\sigma}(\mathbf{r}, t) = i\frac{\partial}{\partial t} \psi_{i\sigma}(\mathbf{r}, t) \tag{29}$$

Here, the wave functions $\psi_i(\mathbf{r}, t)$ and $v_{eff}[\rho](\mathbf{r}, t)$ explicitly depend on time, whereas,

$$v_{eff}[\rho](\mathbf{r}, t) = \sum_a v_{ion}(\mathbf{r} - \mathbf{R}_a) + \int \frac{\rho[\mathbf{r}', t]}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + v_{xc}[\rho](\mathbf{r}, t) \,. \tag{30}$$

Using the adiabatic approximation (AA), (which is local in time)

$$
\begin{aligned}
v_{xc}[\rho](\mathbf{r}, t) &\approx \frac{\delta E_{xc}[\rho]}{\delta \rho(\mathbf{r})} \\
\frac{\delta v_{xc}[\rho](\mathbf{r}, t)}{\delta \rho(\mathbf{r}', t)} &\approx \delta(t - t') \frac{\delta^2 E_{xc}[\rho]}{\delta \rho(\mathbf{r}) \delta \rho(\mathbf{r}')} \,,
\end{aligned}
\tag{31}
$$

and using the LDA,

$$E_{xc}[\rho] = \int \rho(\mathbf{r}) \epsilon_{xc}(\rho(\mathbf{r})) d\mathbf{r} \,. \tag{32}$$

The method that we wish to use here is Casida's approach [48]. This section explains the deriving equations of linear-response (LR) TD-LDA method.

The time-dependent perturbation to the external potential can be written as,

$$\delta v_{eff}[\rho](\mathbf{r}, t) = \delta v_{appl}(\mathbf{r}, t) + \delta v_{SCF}[\rho](\mathbf{r}, t) \tag{33}$$

where,

$$v_{SCF}[\rho](\mathbf{r},t) = \int \frac{\rho(\mathbf{r}',t)}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}' + v_{xc}[\rho](\mathbf{r}',t) \tag{34}$$

The LR of the density matrix (DM) can be written in terms of generalised susceptibility χ as below,

$$\delta\mathbf{P}_{ij\sigma}(\omega) = \sum_{kl\tau} \chi_{ij\sigma,kl\tau}(\omega)\delta v_{kl\tau}^{eff}(\omega) \tag{35}$$

After some mathematical steps, one can end-up with the sum-over-states (SOS) representation of χ,

$$\chi_{ij\sigma,kl\tau}(\omega) = \delta_{ik}\delta_{jl}\delta_{\sigma\tau}\frac{\lambda_{lk\tau}}{\omega-(\omega_{lk\tau})} \tag{36}$$

where $\lambda_{lk\tau} = n_{l\tau} - n_{k\tau}$ the difference in occupation numbers and $\omega_{lk\tau} = \epsilon_{k\tau} - \epsilon_{l\tau}$ the difference between the eigenvalues of lth and kth states. In the basis of Kohn-Sham orbitals$\{\psi_{i\sigma}\}$, we can re-write the LR-DM equation as,

$$\delta\mathbf{P}_{ij\sigma}(\omega) = \frac{\lambda_{ji\sigma}}{\omega-\omega_{ji\sigma}}\left[\delta v_{ij\sigma}^{appl}(\omega) + \delta v_{ij\sigma}^{SCF}(\omega)\right] \tag{37}$$

Now the term δv^{SCF} is complicated because it itself depends on the response of the DM.

$$\delta v_{ij\sigma}^{SCF}(\omega) = \sum \mathbf{K}_{ij\sigma,kl\tau}\delta\mathbf{P}_{kl\tau}(\omega) \tag{38}$$

Where,

$$\mathbf{K}_{ij\sigma,kl\tau} = \frac{\partial v_{ij\sigma}^{SCF}}{\partial \mathbf{P}_{kl\tau}} \tag{39}$$

whose integral form is,

$$\mathbf{K}_{ij\sigma,kl\tau} = \int\int \psi_{i\sigma}^*(\mathbf{r})\psi_{j\sigma}(\mathbf{r})\left[\frac{1}{|\mathbf{r}-\mathbf{r}'|} + \frac{\delta^2 E_{xc}[\rho]}{\delta\rho_\sigma(\mathbf{r})\delta\rho_\tau(\mathbf{r}')}\right]\psi_{k\tau}(\mathbf{r}')\psi_{l\tau}^*(\mathbf{r}')d\mathbf{r}d\mathbf{r}' \tag{40}$$

If the response is due to a real spin independent external perturbation, δv^{appl}, then we can restrict ourselves to the real density response and the coupling matrix is symmetric.

After some algebra, the real parts of the DM elements $\Re\delta\mathbf{P}(\omega)$ can be given as,

$$\sum_{kl\tau}\left[\frac{\delta_{ik},\delta_{jl}\delta_{\sigma\tau}}{\lambda_{kl\tau}\omega_{kl\tau}}(\omega^2-\omega_{kl\tau}^2) - 2\mathbf{K}_{ij\sigma,kl\tau}\right]\Re(\delta\mathbf{P}_{kl\tau})(\omega) = \delta v_{ij\sigma}^{appl}(\omega) \tag{41}$$

Here the real part of $\Re\delta\mathbf{P}_\sigma(\omega)$ means the Fourier transform of the real part of $\Re\delta\mathbf{P}_\sigma(t)$. Thus the real part of the first-order DM obtained from the solution of the above linear equations gives access to the frequency-dependent polarizabilities. This leads to the following eigenvalue equation from which the excitation energies and oscillator strengths can be obtained.

$$\mathbf{\Omega}\vec{F}_I = \omega_I^2\vec{F}_I\,, \tag{42}$$

where,

$$\mathbf{\Omega}_{ij\sigma,kl\tau} = \delta_{ik}\delta_{jl}\delta_{\sigma\tau}\omega_{kl\tau}^2 + 2\sqrt{\lambda_{ij\sigma}\omega_{ij\sigma}}\mathbf{K}_{ij\sigma,kl\tau}\sqrt{\lambda_{kl\tau}\omega_{kl\tau}} \tag{43}$$

Here, the desired excitation energies are equal to $\omega_{\mathbf{I}}$ and the oscillator strengths f_I are obtained from the eigenvectors $\vec{F}_I$. The frequency-dependent polarizability is directly realted to vertical excitation energies, oscillator strength and transition dipole moments μ_I,

$$\alpha(\omega) = \sum_I \frac{f_I}{\omega_I{}^2 - \omega^2} = \frac{2}{3} \sum_I \frac{\omega_{\mathbf{I}}\mu_I^2}{\omega_{\mathbf{I}}{}^2 - \omega^2} \tag{44}$$

5. Krylov Space Methods

The methods described in this article involve solving very large eigenvalue problems. One of these is the matrix form of the Kohn-Sham orbital equation Eq. (24) while the other is the LR-TD-DFT equation Eq. (42). The first is very large because the wavelet basis set is very large while the other is very large because it is of the order of the number of unoccupied orbitals times the number of unoccupied orbitals on each side. It is important to realize that special methods must be and are being used to solve these very large eigenvalue problems. In particular, BIGDFT make use of the block Davidson variant of the Krylov space method to solve the Kohn-Sham equation while BIGDFT and most other codes solving the LR-TD-DFT equation Eq. (42) also make use of the the block Davidson method. Krylov methods and the block Davidson method are briefly described in this section.

Krylov space methods may be traced back to a paper in the early 1930s written by the Russian mathematician Alexei Nikolaevich Krylov. The main idea is that to solve a matrix problem involving a matrix $\mathbf{A}$, it is frequently never actually necessary to construct the matrix $\mathbf{A}$ because iterative solutions only require a reasonable first guess followed by repeated action of the *operator* $\hat{A}$ on a vector. A number of such methods are known with Lanczos diagonalization and the discrete inversion in the iterative subspace (DIIS) [63] as particlarly well-known examples in theoretical chemical physics. Given a vector $\vec{x}$, the Krylov space of order r is given by,

$$\mathcal{K}_r(\mathbf{A}, \vec{x}) = \text{span} \left\{ \vec{x}, \mathbf{A}\vec{x}, \mathbf{A}^2\vec{x}, \cdots, \mathbf{A}^r\vec{x} \right\} . \tag{45}$$

The Davidson diagonalization method [64] for solving the matrix eigenvalue problem

$$\mathbf{A}\vec{x} = a\vec{x}, \tag{46}$$

is deceptively simple. Suppose that we want the lowest eigenvalue and eigenvector and we have an intial guess vector, $\vec{x}^{(0)}$. Then we can always write,

$$\vec{x} = \vec{x}^{(0)} + \delta\vec{x}, \tag{47}$$

is the component of the exact solution which is orthogonal to the intial gues vector. Simple algebra then gives a formula highly reminiscent of Rayleigh-Schrödinger perturbation theory but exact,

$$\delta\vec{x} = [\mathbf{Q}(a\mathbf{1} - \mathbf{A})\mathbf{Q}]^{-1} (\mathbf{A} - a\mathbf{1}) \vec{x}^{(0)}, \tag{48}$$

where,

$$\mathbf{Q} = \mathbf{1} - \vec{x}^{(0)}\vec{x}^{(0)\dagger}, \tag{49}$$

projects onto the subspace orthogonal to the guess vector. Solving Eq. (48) requires us to overcome two difficulties. The first is that we need a guess for the eigenvalue a, but this is easily remedied by taking $a^{(0)} = \vec{x}^{(0)\dagger} \mathbf{A} \vec{x}^{(0)} / \vec{x}^{(0)\dagger} \vec{x}^{(0)}$ and then iterating. The problem greater problem is to invert the matrix $[\mathbf{Q}(a\mathbf{1} - \mathbf{A})\mathbf{Q}]$. It actually turns out that a highly accurate inversion is not really needed (and actually can cause some problems.) Instead it is better just to replace this difficult inversion with,

$$\delta\vec{x} = (a\mathbf{1} - \mathbf{D})^{-1} (\mathbf{A} - a\mathbf{1}) \vec{x}^{(0)}, \tag{50}$$

where $\mathbf{D}$ is some diagonal matrix, hence easy to invert. Orthogonalizing $\delta\vec{x}$ to $\vec{x}^{(0)}$ and normalizing produces $\vec{x}^{(1)}$, which is the next basis vector in our Krylov space. Setting up and diagonalizing the 2×2 matrix of $\mathbf{A}$ in this basis and taking the lowest eigenvalue gives us the next estimate $a^{(1)}$. If application of Eq. (48) is close to zero then we have solved the eigenvalue problem Eq. (46), otherwise we have a new $\delta\vec{x}$ with which to generate $\vec{x}^{(2)}$ and so on and so forth until convergence. The block Davidson method [65] extends the Davidson method to the lowest several eigenvalues and eigenvectors.

Davidson diagonalization works well when started from a reasonably good initial guess, otherwise the Lanczos method may be advantageous. One of the most recent incarnations of the Lanczos method is the Liouville-Lanczos method for solving the LR-TD-DFT problem [66, 67, 68, 69].

6.   Numerical Implementation of DFT in BIGDFT

Computational physics/chemistry is the transformation and implementation of scientific theory into efficient algorithms which requires both theoretical and experimental skill. The transformation of a new theory into an efficient algorithm requires understanding of programming concepts, mathematical and physical intuition and theoretical insight, whereas the production of the computer code is much like experimentation, requiring debugging, testing and organisation to yield a highly efficient product. It is also an adaptation of new scientific theory into computer code exploiting the advances in compiler, programming language and hardware technology. The aim is to afford an algorithm to enable efficient computation, portability of code, ease of adaptability and to document the science. To afford such an algorithm requires an intuitive understanding of the physics to be implemented, much experimentation with optimisation and debugging of the developing code, a suitable choice of programming language, as well as a basic overview of the nature of the platforms for which the code is intended.

The Kohn-Sham scheme of DFT greatly reduces the complexity of ground state electronic structure calculations by recasting the many-body problem into a (self-consistent) single-particle problem. For real atomistic systems, however, the KS equations are still difficult to solve and further approximate techniques are required. In general it is important, though, that these approximations can be controlled in such a way that the associated error does not exceed the error introduced by the xc-functional. While DFT accounts for approximately 90% of all quantum chemical calculations being performed, the sometimes unpredictable nature of results and the inability to systematically improve the quality of calculation may mean that a place for the conventional correlated techniques remains in the

quantum chemist's tool kit. In this work the detailed description of DFT program BIGDFT has been given. BIGDFT [57] has been developed as an European project (FP6-NEST) from 2005 to 2008, and is a wavelet-based pseudopotential implementation of DFT and TD-DFT. For complimentary purpose, Gaussian based quantum chemistry DFT code DE-MON2K [70] is also used but we are not going to discuss the numerical implementation of DEMON2K here and we restrict ourselves to use DEMON2K for revalidating our recent implementation of LR-TD-DFT in BIGDFT. However in the following sections, we are going to recast how the fundamental computational operations were performed in BIGDFT.

6.1. Daubechies Wavelets

Before embarking on our own endeavours, we should make some reference to related work. First, it should be acknowledged that a considerable amount of work has been done already in pursuit of a wavelets in the electronic structure calculations [71, 72, 73, 74, 75, 76]. The object of using wavelets as basis set is to associate an expansion coefficient to each of the piece-wise wavelets. The expansion coefficients are free to vary from one wavelet function to the next. This feature enables wavelets as highly localised continuous functions of a fractal nature that have finite supports. The Daubechies wavelets have no available analytic forms, and they are not readily available in sampled versions. They are defined effectively by the associated dilation coefficients. These express a wavelet in high resolution and a scaling function in the low resolution–which has the same width and which stretches to zero–as a linear combination of the more densely packed and less dispersed scaling functions that form a basis for the two resolution level in combination.

The fact that the Daubechies wavelets are known only via their dilation coefficients is no implediment to the discrete wavelet transform. This transform generates the expansion co-efficients associated with the wavelet decomposition of a data sequence. In this perspective, the dilation coefficients of the wavelets and of the associated scaling functions are nothing but the coefficients of a pair of quadrature mirror filters that are applied succesively.

As described above Daubechies family consists of two fundamental functions: the scaling function $\phi(x)$ and the wavelet $\varphi(x)$ (see Fig. 4.) The full basis set can be obtained from all translations by a certain grid spacing h of the scaling and wavelet functions centered at the origin. These functions satisfy the fundamental defining (refinement) equations,

$$
\begin{aligned}
\phi(x) &= \sqrt{2} \sum_{j=1-m}^{m} h_j \phi(2x - j), \\
\varphi(x) &= \sqrt{2} \sum_{j=1-m}^{m} g_j \phi(2x - j).
\end{aligned}
\tag{51}
$$

which relate the basis functions on a grid with spacing h and another one with spacing $h/2$. The coefficients, h_j and g_j, consitute the so-called "filters" which define the wavelet family of order m. These coefficients satisfy the relations, $\sum_j h_j = 1$ and $g_j = (-1)^j h_{-j+1}$. Eq. (51) is very important since it means that a scaling-function basis defined over a fine grid of spacing $h/2$ may be replaced by combining a scaling-function basis over a coarse grid of spacing h with a wavelet basis defined over the fine grid of spacing $h/2$. This then gives us the liberty to begin with a coarse description in terms of scaling functions and

then add wavelets only where a more refined description is needed. In principle the refined wavelet description may be further refined by adding higher-order wavelets where needed. However in BIGDFT we restricted ourselves to just two levels: coarse and fine associated respectively with scaling functions and wavelets.

For a three-dimensional description, the simplest basis set is obtained by a tensor product of one-dimensional basis functions. For a two resolution level description, the coarse degrees of freedom are expanded by a single three dimensional function, $\phi^0_{i_1,i_2,i_3}(\mathbf{r})$, while the fine degrees of freedom can be expressed by adding another seven basis functions, $\phi^\nu_{j_1,j_2,j_3}(\mathbf{r})$, which include tensor products with one-dimensional wavelet functions. Thus,
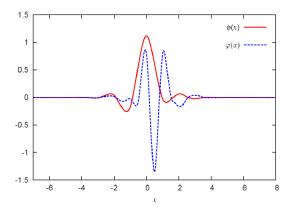

Figure 4. Daubechies scaling function $\phi(x)$ and wavelet $\varphi(x)$ of order 16.

the Kohn-Sham wave function $\psi(\mathbf{r})$ is of the form

$$\psi(\mathbf{r}) = \sum_{i_1,i_2,i_3} c^0_{i_1,i_2,i_3} \phi^0_{i_1,i_2,i_3}(\mathbf{r}) + \sum_{j_1,j_2,j_3} \sum_{\nu_1}^{7} c^\nu_{j_1,j_2,j_3} \phi^\nu_{j_1,j_2,j_3}(\mathbf{r}). \tag{52}$$

The sum over i_1,i_2,i_3 runs over all the grid points contained in the low-resolution regions and the sum over j_1,j_2,j_3 runs over all the points contained in the (generally smaller) high resolution regions. Each wave function is then described by a set of coefficients $\{c^\nu_{j_1,j_2,j_3}\}, \nu = 0, ..., 7$. Only the nonzero scaling function and wavelet coefficients are stored. The data is thus compressed. The basis set being orthogonal, several operations such as scalar products among different orbitals and between orbitals and the projectors of the nonlocal pseudopotential can be directly carried out in this compressed form.

6.2. Treatment of kinetic energy

The matrix elements of the kinetic energy operator among the basis functions of our mixed representation (i.e., scaling functions with scaling functions, scaling function with wavelets and wavelets with wavelets) can be calculated analytically [77]. For simplicity, let us illustrate the application of the kinetic energy operator onto a wavefunction ψ that is only expressed in terms of scaling functions.

$$\psi(x,y,z) = \sum_{i_1,i_2,i_3} s_{i_1,i_2,i_3} \phi(x/h - i_1)\,\phi(y/h - i_2)\,\phi(z/h - i_3)$$

The result of the application of the kinetic energy operator on this wavefunction, projected to the original scaling function space, has the expansion coefficients

$$\hat{s}_{i_1,i_2,i_3} = -\frac{1}{2h^3} \int \phi(x/h - i_1)\, \phi(y/h - i_2)\, \phi(z/h - i_3) \times$$
$$\times \nabla^2 \psi(x,y,z) \mathrm{dxdydz} \ .$$

Analytically the coefficients s_{i_1,i_2,i_3} and $\hat{s}_{i_1,i_2,i_3}$ are related by a convolution

$$\hat{s}_{i_1,i_2,i_3} = \frac{1}{2} \sum_{j_1,j_2,j_3} K_{i_1-j_1,i_2-j_2,i_3-j_3} s_{j_1,j_2,j_3} \tag{53}$$

where

$$K_{i_1,i_2,i_3} = T_{i_1} T_{i_2} T_{i_3}, \tag{54}$$

where the coefficients T_i can be calculated analytically via an eigenvalue equation:

$$
\begin{aligned}
T_i &= \int \phi(x) \frac{\partial^2}{\partial x^2} \phi(x - i) dx \\
&= \sum_{\nu,\mu} 2 h_\nu h_\mu \int \phi(2x - \nu) \frac{\partial^2}{\partial x^2} \phi(2x - 2i - \mu) dx \\
&= \sum_{\nu,\mu} 2 h_\nu h_\mu 2^{2-1} \int \phi(y - \nu) \frac{\partial^2}{\partial y^2} \phi(y - 2i - \mu) dy \\
&= \sum_{\nu,\mu} h_\nu h_\mu 2^2 \int \phi(y) \frac{\partial^2}{\partial y^l} \phi(y - 2i - \mu + \nu) dy \\
&= \sum_{\nu,\mu} h_\nu h_\mu\, 2^2\, T_{2i-\nu+\mu}
\end{aligned}
$$

Using the refinement equation Eq. (51), the values of the T_i can be calculated analytically, from a suitable eigenvector of a matrix derived from the wavelet filters [77]. For this reason the expression of the kinetic energy operator is *exact* in a given Daubechies basis.

Since the 3-dimensional kinetic energy filter K_{i_1,i_2,i_3} is a product of three one-dimensional filters Eq. (54) the convolution in Eq. (53) can be evaluated with $3N_1 N_2 N_3 L$ operations for a three-dimensional grid of $N_1 N_2 N_3$ grid points. L is the length of the one-dimensional filter which is 29 for our Daubechies family. The kinetic energy can thus be evaluated with linear scaling with respect to the number of nonvanishing expansion co-efficients of the wavefunction. This statement remains true for a mixed scaling function-wavelet basis where we have both nonvanishing s and d coefficients and for the case where the low and high resolution regions cover only parts of the cube of $N_1 N_2 N_3$ grid points.

6.3.   Treatment of local potential energy

In spite of the striking advantages of Daubechies wavelets the initial exploration of this basis set [78] did not lead to any algorithm that would be useful for practical electronic structure calculations. This was due to the fact that an accurate evaluation of the local potential energy is difficult in a Daubechies wavelet basis.

By definition, the local potential $v(\mathbf{r})$ can be easily known on the nodes of the uniform grid of the simulation box. Approximating a potential energy matrix element $v_{i,j,k;i',j',k'}$

$$v_{i,j,k;i',j',k'} = \int d\mathbf{r}\phi_{i',j',k'}(\mathbf{r})v(\mathbf{r})\phi_{i,j,k}(\mathbf{r})$$

by

$$v_{i,j,k;i',j',k'} \approx \sum_{l,m,n} \phi_{i',j',k'}(\mathbf{r}_{l,m,n})v(\mathbf{r}_{l,m,n})\phi_{i,j,k}(\mathbf{r}_{l,m,n})$$

gives an extremely slow convergence rate with respect to the number of grid points used to approximate the integral because a single scaling function is not very smooth, i.e., it has a rather low number of continuous derivatives. A. Neelov and S. Goedecker [79] have shown that one should not try to approximate a single matrix element as accurately as possible but that one should try instead to approximate directly the expectation value of the local potential. The reason for this strategy is that the wavefunction expressed in the Daubechy basis is smoother than a single Daubechies basis function. A single Daubechies scaling function of order 16 (i.e., the corresponding wavelet has 8 vanishing moments) has only 2 continuous derivatives. More precisely its index of Hölder continuity is about 2.7 and the Sobolev space regularity with respect to $p = 2$ is about 2.91 [80]. A single Daubechies scaling function of order 16 has only 4 continuous derivatives. By suitable linear combinations of Daubechies 16 one can however exactly represent polynomials up to degree 7, i.e., functions that have 7 non-vanishing continuous derivatives. The discontinuities get thus canceled by taking suitable linear combinations. Since we use pseudopotentials, our exact wavefunctions are analytic and can locally be represented by a Taylor series. We are thus approximating functions that are approximately polynomials of order 7 and the discontinuities nearly cancel.

Instead of calculating the exact matrix elements we therefore use matrix elements with respect to a smoothed version $\tilde{\phi}$ of the Daubechies scaling functions.

$$\begin{aligned} v_{i,j,k;i',j',k'} &\approx \sum_{l,m,n} \tilde{\phi}_{i',j',k'}(\mathbf{r}_{l,m,n})v(\mathbf{r}_{l,m,n})\tilde{\phi}_{i,j,k}(\mathbf{r}_{l,m,n}) \\ &= \sum_{l,m,n} \tilde{\phi}_{0,0,0}(\mathbf{r}_{l-i',m-j',n-k'})V(\mathbf{r}_{l,m,n})\tilde{\phi}_{0,0,0}(\mathbf{r}_{l-i,m-j,n-k}) \quad (55) \end{aligned}$$

where the smoothed wave function is defined by

$$\tilde{\phi}_{0,0,0}(\mathbf{r}_{l,m,n}) = \omega_l\omega_m\omega_n$$

and ω_l is the "magic filter". Even though Eq. (55) is not a particulary good approximation for a single matrix element it gives an excellent approximation for the expectation values of the local potential energy

$$\int dx \int dy \int dz\psi(x,y,z)v(x,y,z)\psi(x,y,z)$$

and also for matrix elements between different wavefunctions

$$\int dx \int dy \int dz\psi_i(x,y,z)v(x,y,z)\psi_j(x,y,z)$$

in case they are needed. Because of this remarkable achievement of the filter ω we call it the magic filter.

Following the same guidelines as the kinetic energy filters, the smoothed real space values $\tilde{\psi}_{i,j,k}$ of a wavefunction ψ are calculated by performing a product of three one-dimensional convolutions with the magic filters along the x, y and z directions. For the scaling function part of the wavefunction the corresponding formula is

$$\tilde{\psi}_{i_1,i_2,i_3} = \sum_{j_1,j_2,j_3} s_{j_1,j_2,j_3} v^{(1)}_{i_1-2j_1} v^{(1)}_{i_2-2j_2} v^{(1)}_{i_3-2j_3}$$

where $v_i^{(1)}$ is the filter that maps a scaling function on a double resolution grid. Similar convolutions are needed for the wavelet part. The calculation is thus similar to the treatment of the Laplacian in the kinetic energy.

Once we have calculated $\tilde{\psi}_{i,j,k}$ the approximate expectation value ϵ_V of the local potential v for a wavefunction ψ is obtained by simple summation on the double resolution real space grid:

$$\epsilon_v = \sum_{j_1,j_2,j_3} \tilde{\psi}_{j_1,j_2,j_3} v_{j_1,j_2,j_3} \tilde{\psi}_{j_1,j_2,j_3}$$

6.4. Treatment of the non-local pseudopotential

The energy contributions from the non-local pseudopotential have for each angular moment l the form

$$\sum_{i,j} \langle \psi | p_i \rangle h_{ij} \langle p_j | \psi \rangle$$

where $|p_i\rangle$ is a pseudopotential projector. Once applying the Hamiltonian operator, the application of one projector on the wavefunctions requires the calculation of

$$|\psi\rangle \to |\psi\rangle + \sum_{i,j} |p_i\rangle h_{ij} \langle p_j | \psi \rangle \ .$$

If we use for the projectors the representation of Eq. (52) (i.e., the same as for the wavefunctions) both operations are trivial to perform. Because of the orthogonality of the basis set we just have to calculate scalar products among the coefficient vectors and to update the wavefunctions. The scaling function and wavelet expansion coefficients for the projectors are given by [81]

$$\int p(\mathbf{r}) \, \phi_{i_1,i_2,i_3}(\mathbf{r}) \mathrm{d}\mathbf{r} \ , \quad \int \mathrm{p}(\mathbf{r}) \, \varphi^{\nu}_{i_1,i_2,i_3}(\mathbf{r}) \mathrm{d}\mathbf{r} \ . \tag{56}$$

The GTH-HGH pseudopotentials [82, 83] have projectors which are written in terms of gaussians times polynomials. This form of projectors is particularly convenient to be expanded in the Daubechies basis. In other terms, since the general form of the projector is

$$\langle \mathbf{r} | p \rangle = e^{-cr^2} x^{\ell_x} y^{\ell_y} z^{\ell_z} \ ,$$

the 3-dimensional integrals can be calculated easily since they can be factorized into a product of 3 one-dimensional integrals.

$$\int \langle \mathbf{r}|p\rangle \phi_{i_1,i_2,i_3}(\mathbf{r})\mathrm{d}\mathbf{r} = W_{i_1}(c,\ell_x)W_{i_2}(c,\ell_y)W_{i_3}(c,\ell_x)\,, \tag{57}$$

$$W_j(c,\ell) = \int_{-\infty}^{+\infty} e^{-ct^2} t^\ell \phi(t/h - j)\mathrm{dt} \tag{58}$$

The one-dimensional integrals are calculated in the following way. We first calculate the scaling function expansion coefficients for scaling functions on a one-dimensional grid that is 16 times denser. The integration on this dense grid is done by the well-known quadrature introduced in [84], that coincides with the magic filter [79]. This integration scheme based on the magic filter has a convergence rate of h^{16} and we gain therefore a factor of 16^{16} in accuracy by going to a denser grid. This means that the expansion coefficients are for reasonable grid spacings h accurate to machine precision. After having obtained the expansion coefficients with respect to the fine scaling functions we obtain the expansion coefficients with respect to the scaling functions and wavelets on the required resolution level by one-dimensional fast wavelet transformations. No accuracy is lost in the wavelet transforms and our representation of the projectors is therefore typically accurate to nearly machine precision. In order to treat with the same advantages other pseudopotentials which are not given under the form of gaussians it would be necessary to approximate them by a small number of gaussians.

6.5. The Poisson operator

Solving the Poisson equation for an arbitrary charge distribution is a non-trivial task, and is of major importance in many field of science, especially in the field of computational chemistry. A huge effort has been put into making efficient Poisson solvers, and usual real-space approaches includes finite difference (FD) and finite element (FE) methods. FD is a grid-based method, which is solving the equations iteratively on a discrete grid of pointvalues, while FE is expanding the solution in a basis set, usually by dividing space into cubic cells and allocate a polynomial basis to each cell.

It is well-known fact that the electronic density in molecular systems is rapidly varying in the vicinity of the atomic nuclei, and a usual problem with real-space methods is that an accurate treatment of the system requires high resolution of gridpoints (FD) or cells (FE) in the nuclear regions. Keeping this high resolution uniformly throughout space would yield unnecessary high accuracy in the interatomic regions, and the solution of the Poisson equation for molecular systems is demanding a *multiresolution* framework in order to achieve numerical efficiency. This chapter is concerned with a way of doing DFT and TD-DFT calculations, one where the multiresolution character is inherent in the theory, namely using wavelet bases.

In order to evaluate the Hartree potential, we need to rewrite the standard Poisson equation to an integral form. The equation, in its differential form, is given as

$$\nabla^2 v(\mathbf{x}) = 4\pi\rho(\mathbf{x})\,, \tag{59}$$

where $\rho(\mathbf{x})$ is the known (charge) distribution, and $v(\mathbf{x})$ is the unknown (electrostatic) potential. It is a standard textbook procedure to show that the solution can be written as the integral

$$v(\mathbf{x}) = \int G(\mathbf{x}, \mathbf{y}) \rho(\mathbf{y}) d\mathbf{y} \, , \tag{60}$$

where $G(\mathbf{x}, \mathbf{y})$ is the Green's function which is the solution to the *fundamental* equation with *homogeneous* (Dirichlet) boundary conditions

$$\begin{aligned} \nabla^2 G(\mathbf{x}, \mathbf{y}) &= \delta(\mathbf{x} - \mathbf{y}) \\ G(\mathbf{x}, \mathbf{y}) &= 0 \, , \mathbf{x} \in \text{boundary} \end{aligned} \tag{61}$$

This equation can be solved analytically and the Green's function is given (in three dimensions) simply as

$$G(\mathbf{x}, \mathbf{y}) = \frac{1}{||\mathbf{x}, \mathbf{y}||} \, , \tag{62}$$

This is the well-known potential arising from a point charge located in the position $\mathbf{y}$, which is exactly what Eq. (61) describes.

6.6. Numerical separation of the kernel

The Green's function kernel as it is given in Eq. (62) is not separable in the cartesian coordinates. However, since we are working with finite precision we can get by with an *approximate* kernel as long as the error introduced with this approximation is less than our overall accuracy criterion. If we are able to obtain such a *numerical* separation of the kernel, the operator can be applied in one direction at the time, allowing us to use the expressions derived above for one-dimentional integral operators to solve the three-dimensional Poisson equation. This is of great importance because it reduces the scaling behavior to become linear in the dimension of the system.

The Poisson kernel can be made separable by expanding it as a sum of Gaussian functions, specifically

$$\frac{1}{r} \simeq \sum_{k=1}^{M_\epsilon} \omega_k e^{-p_k r^2} \, . \tag{63}$$

where ω_k and p_k are parameters that needs to be determined, and the number of terms M_ϵ, called the separation rank, depends on the accuracy requirement and on what interval this expression needs to be valid. Details of how to obtain this expression can be found in [71, 72], and will not be treated here, but it should be mentioned that the separation rank is usually in the order of 100, e.g, it requires $M_\epsilon = 89$ to reproduce $\frac{1}{r}$ on the interval $[10^{-9}, 1]$ in three dimensions with error less than $\epsilon = 10^{-8}$.

Finally, figure 5 summarize this complete section into a flow-chart type diagram. This kind of explanation is necessary for begineers because there are different functions used for the different operations in BIGDFT. As one can see from the figure, The KS wavefunctions $|\psi\rangle$ are expressed in terms of Daubechies wavelets and the projection of Hamiltonian $V_{nl}|\psi\rangle$ and of pseudopotential operators $|H\psi\rangle$ also expressed using Daubechies wavelets. The rest of the operations such as kinetic energy $|\nabla^2 \psi\rangle$, potential energy operator $V(x)\psi(x)$, and

the local densities $\rho(x)$ are all expressed using interpolating scaling functions, in which the Hartree $V_H(x)$, local part potential energy $V_{loc}(x)$ and xc operations $V_{xc}(x)$ were performed in real space. The interconnecting lines between different operations represents the transformation between Daubechies wavelets-to-ISFs or the transformation of real space-to-fourier space representation.
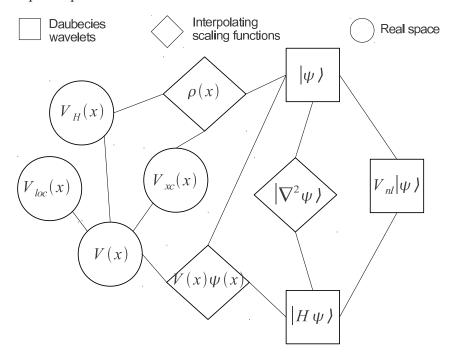

Figure 5. Operations performed in BIGDFT

7.   BIGDFT **and TD-DFT**

We want to solve Casida's equation [48],

$$\left[\left(\begin{array}{cc} \mathbf{A}(\omega) & \mathbf{B}(\omega) \\ \mathbf{B}^*(\omega) & \mathbf{A}^*(\omega) \end{array} \right) - \omega \left(\begin{array}{cc} 1 & 0 \\ 0 & -1 \end{array} \right) \right] \left(\begin{array}{c} \mathbf{X} \\ \mathbf{Y} \end{array} \right) = 0 , \qquad (64)$$

where $\mathbf{X}$ and $\mathbf{Y}$ represents the pseudo eigenvectors; the matrices $\mathbf{A}$ and $\mathbf{B}$ are defined as

$$\mathbf{A}_{ai\sigma,bj\tau} = \delta_{ab}\delta_{ij}\delta_{\sigma\tau}(\epsilon_a - \epsilon_i) + \mathbf{K}_{ai\sigma,bj\tau}(\omega) , \qquad (65)$$

and,

$$\mathbf{B}_{ai\sigma,bj\tau} = \mathbf{K}_{ai\sigma,jb\tau}(\omega) , \qquad (66)$$

in which the integral form of the coupling matrix $\mathbf{K}$ is given by,

$$\mathbf{K}_{pq\sigma,rs\tau} = \int \int \Psi_{p\sigma}^*(\vec{r})\Psi_{q\sigma}(\vec{r})$$

$$\left[\frac{1}{|\vec{r} - \vec{r'}|} + \frac{\partial^2 E_{xc}[\rho]}{\partial \rho_\sigma(\vec{r})\partial \rho_\tau(\vec{r'})} \right] \Psi_{r\tau}(\vec{r'})\Psi_{s\tau}^*(\vec{r'})d\vec{r}d\vec{r'} . \qquad (67)$$

The universal adiabatic approximation is applied to Eq. (67) to remove the frequency dependence of the kernel.

The electronic transitions occur with an infinitesimal perturbation obtains the above described non-Hermitian eigenvalue Eq. (64). Where the response is due to a real spin independent external perturbation, and the actual response is described as the real density response. However, an unitary transformation is necessary to convert Eq. (64) into the real eiganvalue problem. In Eq. (64), all occupied-occupied and virtual-virtual element contributions are zero whereas only the elements that are from virtual-occupied and occupied-virtual parts are taken into account. Moreover if we only restricted to virtual-occupied elements and neglecting the occupied-virtual elements of Eq. (64) leads to a Hermitian eigenvalue equation of the dimension one-half of that TD-DFT working equation is said to be Tamm-Dancoff approximation (TDA) and it is written as,

$$\mathbf{AX} = \omega \mathbf{X}\,, \tag{68}$$

where $\mathbf{A}$ is as same as in Eq. (42). The matrix $\mathbf{A}$ is just restricted to number of single excitations.

However, the explicit form of Eq. (68) is,

$$\Omega(\omega)\vec{F}_I = \omega^2 \vec{F}_I\,, \tag{69}$$

where

$$\begin{aligned}
\Omega_{i a\sigma, jb\tau} &= \delta_{ia}\delta_{jb}\delta_{\sigma\tau}(\epsilon_{a\sigma} - \epsilon_{i\sigma})^2 + \\
&\quad 2\sqrt{(\epsilon_{a\sigma} - \epsilon_{i\sigma})}K_{i a\sigma, jb\tau}\sqrt{(\epsilon_{a\sigma} - \epsilon_{i\sigma})}\,,
\end{aligned} \tag{70}$$

where $\epsilon_{i\sigma} - \epsilon_{a\sigma}$ is the energy eigenvalue differences of ith and a^{th} states. Solving Eqs. (69) yields TD-DFT excitation energies ω and $\vec{F}_I$'s are the corresponding oscillator strengths which are defined from the transition dipole moments.

7.1. Calculation of Coupling Matrix

We are now in a position to understand the construction of the coupling matrix Eq. (67) in our implementation of TD-DFT in BIGDFT, which we split into the Hartree andxc parts,

$$K_{ai\sigma, bj\tau} = K^H_{ai\sigma, bj\tau} + K^{xc}_{aj\sigma, bj\tau}\,. \tag{71}$$

Instead of calculating the Hartree part of coupling matrix directly as,

$$K^H_{ai\sigma, bj\tau} = \int\int \psi^*_{a\sigma}(\mathbf{r})\psi_{i\sigma}(\mathbf{r})\frac{1}{|\mathbf{r} - \mathbf{r}'|}\psi_{b\tau}(\mathbf{r}')\psi^*_{j\tau}(\mathbf{r}')\,d\mathbf{r}d\mathbf{r}'\,, \tag{72}$$

we express the coupling matrix element as,

$$K^H_{ai\sigma, bj\tau} = \int \psi^*_{a\sigma}(\mathbf{r})\psi_{i\sigma}(\mathbf{r})v_{bj\tau}(\mathbf{r})\,d\mathbf{r}\,, \tag{73}$$

where,

$$v_{ai\sigma}(\mathbf{r}) = \int \frac{\rho_{ai\sigma}(\mathbf{r})}{|\mathbf{r} - \mathbf{r}'|}\,d\mathbf{r}'\,, \tag{74}$$

and,

$$\rho_{ai\sigma}(\mathbf{r}) = \psi^*_{a\sigma}(\mathbf{r})\psi_{i\sigma}(\mathbf{r}) \,. \tag{75}$$

The advantage of doing this is that, although $\rho_{ai\sigma}$ and $v_{ai\sigma}$ are neither real physical charge densities nor real physical potentials, they still satisfy the Poisson equation,

$$\nabla^2 v_{ai\sigma}(\mathbf{r}) = -4\pi\rho_{ai\sigma}(\mathbf{r}) \,, \tag{76}$$

and we can make use of whichever of the efficient wavelet-based Poisson solvers already available in BIGDFT, is appropriate for the boundary conditions of our physical problem.

Once the solution of Poisson's equation, $v_{ai\sigma}(\mathbf{r})$, is known, we can then calculate the Hartree part of the kernel according to Eq. (73). Inclusion of the xc kernel is accomplished by evaluating,

$$K_{ai\sigma,bj\tau} = \int M_{ai\sigma}(\mathbf{r})\rho_{bj\tau}(\mathbf{r})\,d\mathbf{r} \,, \tag{77}$$

where,

$$M_{ai\sigma}(\mathbf{r}) = v_{ai\sigma}(\mathbf{r}) + \int \rho_{ai\sigma}(\mathbf{r}')f^{\sigma;\tau}_{xc}(\mathbf{r},\mathbf{r}')\,d\mathbf{r}' \,. \tag{78}$$

We note that $f^{\sigma;\tau}_{xc}(\mathbf{r},\mathbf{r}') = f^{\sigma;\tau}_{xc}(\mathbf{r},\mathbf{r}')\delta(\mathbf{r}-\mathbf{r}')$ for the LDA, so that no integral need actually be carried out in evaluating $M_{ai\sigma}(\mathbf{r})$. The integral in Eq. (77) is, of course, carried out numerically in practice as a discrete summation.

8.   Results

We now wish to illustrate a bit how wavelet calculations work in the BIGDFT program. Comparison will be made against results obtained with the GTO-based program DEMON2K. This work is very similar to our previous work reporting the first implementation of wavelet-based TD-DFT with illustration for N_2 and application to the absorption spectrum of a medium-sized organic molecule of potential biomedical use as a fluorescent probe [45]. Here however we will present new BIGDFT results for a different small molecule, namely carbon monoxide. Though CO is roughly isoelectronic with N_2, CO has the interesting feature of having a low-lying bright state in its absorption spectrum.

8.1.   Computational Details

Calculations were carried out with DEMON2K and BIGDFT with the LDA-optimized bond length of 1.129 Å.

8.1.1.   DEMON2K

DEMON2K resembles a typical GTO-based quantum chemistry program in that all the integrals other than the xc-integrals, can be evaluated analytically. In particular, DEMON2K has the important advantage that it accepts the popular GTO basis sets common in quantum chemistry and so can benefit from the experience in basis set construction of a large community built up over the past 50 years or so. In the following, we have chosen to use the well-known correlation-consistent basis sets for this study [85, 86]. (Note, however,

that the correlation-consistent basis sets used in DEMON2K lack f and g functions but are otherwise exactly the same as the usual ones.) The advantage of using these particular basis sets is that there is a clear hierarchy as to quality.

An exception to the rule that integrals are evaluated analytically in DEMON2K are the xc-integrals (for the xc-energy, xc-potential, and xc-kernel) which are evaluated numerically over a Becke atom-centered grid. This is important because the relative simplicity of evaluating integrals over a grid has allowed the rapid implemenation of new functionals as they were introduced. We made use of the fine fixed grid in our calculations.

As described so far, DEMON2K should have $\mathcal{O}(N^4)$ scaling because of the need to evaluate 4-center integrals. Instead DEMON2K uses a second atom-centered auxiliary GTO basis to expand the charge density. This allows the the elimination of all 4-center integrals so that only 3-center integrals remain for a formal $\mathcal{O}(N^3)$ scaling. In practice, integral prescreening leads to $\mathcal{O}(NM)$ scaling where M is typically between 2 and 3. We made use of the A3 auxiliary basis set from the DEMON2K automated auxiliary basis set library.

All calculations were performed using standard DEMON2K default criteria. The implementation of TD-DFT in DEMON2K is described in Ref. [87]. (The charge density conservation constraint is no longer used in DEMON2K TD-DFT calculations.) Although full TD-LDA calculations are possible with DEMON2K, the TD-LDA calculations reported here all made use of the TDA.

8.1.2. BIGDFT

(a) H_2O in a box (b) H_2O in a box showing fine grid resolution (c) H_2O in a box showing coarse grid resolution
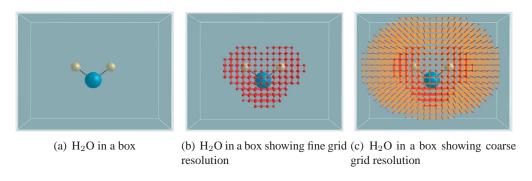
Figure 6. Adaptive grid in BIGDFT (a), (b) and (c)

The main thing to vary in BIGDFT is the grid which is of more profound importance than in DEMON2K because it is the grid which supports the wavelets. Figures 6(a), 6(b), and 6(c) give an idea of what the grid looks like for the small familiar molecule of water. Conceptually the molecule is in a very large box (Fig. 6(a).) A fine grid is placed in the regions of high electron density around the molecule (Fig. 6(b).) A coarse grid is used in a larger region where the electron density varies more slowly (Fig. 6(c).) The BIGDFT grid is characterized by the triple h_g/crmult/frmult. The first number in the triple (h_g) is a real number which specifies the nodes of the grid in atomic units. The second number (the integer-valued crmult) is the coarse grid multiplier. And the third number (the integer-valued frmult) is the fine grid multiplier. Two points must be clearly understood when looking at Figs. 6(a), 6(b), and 6(c). The first is that, while the box may determine the limits

Table 1. Basis set dependence of the HOMO and LUMO energies and of the HOMO-LUMO gap (eV) calculated using DEMON2K.

Basis Set	$-\epsilon_{HOMO}$	$-\epsilon_{LUMO}$	$\Delta\epsilon_{HOMO-LUMO}$
STO-3G	-5.5350	1.2428	4.2922
DZVP	-8.9271	-2.0942	6.8329
TZVP	-9.0287	-2.1902	6.8385
CC-PVDZ	-8.6729	-1.7823	6.8906
CC-PVTZ	-9.0419	-2.1195	6.9224
CC-PVQZ	-9.0944	-2.1971	6.8973
CC-PV5Z	-9.1169	-2.2400	6.8769
CC-PCVDZ	-8.6905	-1.7922	6.8983
CC-PCVQZ	-9.0957	-2.1988	6.8969
CC-PCVTZ	-9.0371	-2.1165	6.9206
CC-PCV5Z	-9.1172	-2.2401	6.8771
AUG-CC-PVDZ	-9.0910	-2.2345	6.8565
AUG-CC-PVQZ	-9.1286	-2.2567	6.8719
AUG-CC-PVTZ	-9.1306	-2.2535	6.8771
AUG-CC-PV5Z	-9.1289	-2.2606	6.8683
AUG-CC-PCVDZ	-9.0987	-2.2371	6.8616
AUG-CC-PCVTZ	-9.1316	-2.2554	6.5776
AUG-CC-PCVQZ	-9.1293	-2.2574	6.8719
AUG-CC-PCV5Z	-9.1291	-2.2607	6.8684

of the grid, the grid does not have the shape of the box and there are no basis functions where there are no grid points. This means that we are not dealing with box boundary conditions, but rather with effective boundary conditions which reflect the shape of the molecule. The other point which is not brought out by our explanation is that the BIGDFT grid is adaptive in the sense that additional fine grid points are added during the calculation as they are needed to maintain and improve numerical precision.

The implementation of TD-DFT in BIGDFT is described in Ref. [45].

8.2. Orbital Energies

Possibly the most remarkable property of wavelets is how rapidly they converge to the basis set limit. Let us illustrate this by comparing highest-occupied molecular orbital (HOMO) and lowest-unoccupied molecular orbital (LUMO) energies calculated with DE-MON2K and BIGDFT. The difference of these two energies is the HOMO-LUMO gap, $\Delta\epsilon_{HOMO-LUMO}$.

Consider first how DEMON2K calculations of $\Delta\epsilon_{HOMO-LUMO}$, evolve as the basis set is improved (Table 1.) Convergence to the true HOMO-LUMO LDA gap is expected

Table 2. Basis set dependence of the HOMO and LUMO energies and of the HOMO-LUMO gap (eV) calculated using BIGDFT.

$h_g{}^a/m^b/n^c$	$-\epsilon_{HOMO}$	$-\epsilon_{LUMO}$	$\Delta\epsilon_{HOMO-LUMO}$
0.4/6/8	-9.0976	-2.1946	6.9029
0.4/7/8	-9.1014	-2.2028	6.8985
0.4/8/8	-9.1017	-2.2044	6.8971
0.4/9/8	-9.1017	-2.2049	6.8967
0.4/10/8	-9.1017	-2.2049	6.8966
0.3/7/8	-9.1022	-2.2056	6.8964
0.3/8/8	-9.1025	-2.2073	6.8950

[a] Grid spacing of the cartesian grid in atomic units.
[b] Coarse grid multiplier (crmult).
[c] Fine grid multiplier (frmult).

with systematic improvement within the series: (i) double zeta plus valence polarization (DZVP) → triple zeta plus valence polarization (TZVP), (ii) augmented correlation-consistent double zeta plus polarization plus diffuse on all atoms (AUG-CC-PCVDZ) → AUG-CC-PCVTZ (triple zeta) → AUG-CC-PCVQZ (quadruple zeta) → AUG-CC- PCV5Z (quintuple zeta), (iii) augmented correlation-consistent valence double zeta plus polarization plus diffuse (AUG-CC-PVDZ) → AUG-CC-PVTZ → AUG-CC-PVQZ → AUG-CC-PV5Z, (iv) correlation-consistent double zeta plus polarization plus tight core (CC-PCVDZ) → CC-PCVTZ → CC-PCVQZ → CC-PCV5Z, and (v) correlation-consistent valence double zeta plus polarization on all atoms (CC-PVDZ) → CC-PVTZ → CC-PVQZ → CC-PV5Z. There is a clear tendency in the correlation-consistent basis sets to tend towards values of -9.13 eV for the HOMO energy, -2.26 eV for the LUMO energy, and 6.87 eV for $\Delta\epsilon_{HOMO-LUMO}$, with adequate convergance achieved with the AUG-CC-PVQZ basis set.

Now let us turn to BigDFT (Table 2). Calculations were done for several different grids, including the high-resolution combination 0.3/8/8 and the low-resolution combination of 0.4/6/8. Remarkably, except for the very lowest quality grid 0.4/6/8, there is essentially no difference between results obtained with the two grids (and even the 0.4/6/8 grid gives nearly converged results.) The results are also quite close to, but not identical to those obtained with the DEMON2K program. The reason for the small differences between the converged results obtained with the two programs is more difficult to trace as it might be due to the auxiliary basis approximation in DEMON2K or to the use of pseudopotentials in BIGDFT or perhaps to still other program differences. The important point is that differences are remarkably small.

8.3. Excitation Energies

Orbital energy differences provide a first estimate for excitation energies. In this case, we would expect to see the HOMO → LUMO excitation at $\Delta\epsilon_{HOMO-LUMO} \approx 6.9$ eV (6.87 eV for DEMON2K and 6.90 eV for BIGDFT.) A better estimate is provided by the two-orbital two-electron model (TOTEM) [48, 88, 89, 90] for the singlet (S) and triplet (T)

Table 3. Comparison of lowest excitation energies of CO (in eV) calculated using BIGDFT and DEMON2K and with experiment.

State	BIGDFT[a]	DEMON2K[b]	Experiment[c]
$1^3\Sigma^-$	9.84	9.85	9.88
$1^3\Delta$	9.17	9.21	9.36
$1^1\Pi$	8.94	8.42	8.51
$1^3\Sigma^+$	8.94	8.54	8.51
$1^3\Pi$	6.47	6.05	6.32

[a] Present work (TD-LDA/TDA) using AUG-CC-PCQZ basis set.
[b] Present work (TD-LDA/TDA) using 0.3/8/8 grid.
[c] Taken from Ref. [91].

transition from orbital i to orbital a,

$$\begin{aligned} \hbar\omega_{i\to a}^T &= \Delta\epsilon_{i\to a} + (ia|f_{xc}^{\alpha,\alpha} - f_{xc}^{\alpha,\beta}|ai) \\ \hbar\omega_{i\to a}^S &= \Delta\epsilon_{i\to a} + (ia|2f_H + f_{xc}^{\alpha,\alpha} + f_{xc}^{\alpha,\beta}|ai) \,, \end{aligned} \tag{79}$$

where

$$\Delta\epsilon_{i\to a} = \epsilon_a - \epsilon_i \,. \tag{80}$$

The TOTEM model often works surprisingly well for small molecules because, unlike the Hartree-Fock approximation which is better adapted to describe electron ionization and attachment, pure DFT Kohn-Sham orbitals are preprepared to describe excitation energies in the sense that the occupied and unoccupied orbitals see the same potential, thus minimizing orbital relaxation effects. Inspection of the sizes and signs of the integrals in Eq. (79) indicates that we should expect,

$$\hbar\omega_{i\to a}^T \le \Delta\epsilon_{i\to a} \le \hbar\omega_{i\to a}^S \,. \tag{81}$$

These is confirmed in Table 3 where the $1^3\Pi$ and $1^1\Pi$ excitations are, respectively, the triplet and singlet states corresponding to the HOMO $\to$ LUMO transition.

Assuming that $f_{xc}^{\alpha,\alpha}$ dominates over $f_{xc}^{\alpha,\beta}$, we may even go a bit further to estimate $(ia|f_H|ai)$ and $(ia|f_{xc}^{\alpha,\alpha}|ai)$ (Fig. 7.) The calculations are show in Table 4. Comparison of $(ia|f_{xc}^{\alpha,\alpha}|ai)^{(1)}$ and $(ia|f_{xc}^{\alpha,\alpha}|ai)^{(2)}$ provides an indication of the quality of the approximation of neglecting the $(ia|f_{xc}^{\alpha,\beta}|ai)$ integral which in this case appears to be excellent. The $(ia|f_H|ai)$ integrals calculated with the two programs are reasonably close. Interestingly the $(ia|f_{xc}^{\alpha,\alpha}|ai)$ disagree by about 0.4 eV which, though small, is not negligible.

Let us now examine the issue of the collapse of the continuum. In Ref. [92], it was shown that the TD-DFT ionization continuum begins at $-\epsilon_{HOMO}$. In exact Kohn-Sham DFT, this should be the ionization potential. However typical approximate density functionals underbind electrons and so lead to an artificially-early on-set of the TD-DFT ionization continuum. This is first illustrated using the DEMON2K program and different basis sets. Indeed Fig. 8 shows that the states above $-\epsilon_{HOMO}$ tend to collapse towards $-\epsilon_{HOMO}$ rather than converging as they should. This is simply because we are trying to describe a continuum which should not be there with a finite basis set. Also seen in the figure is a

Figure 7. Estimation of integrals within the TOTEM model.

Table 4. Estimations of integrals (in eV) within the TOTEM.

Program	DEMON2K	BIGDFT
Input Data		
$1^1\Pi$	8.42	8.94
$\Delta\epsilon_{i\rightarrow a}$	6.87	6.90
$1^3\Pi$	6.05	6.47
Derived Results		
$(ia|f_H|ai)$	1.19	1.24
$(ia|f_{xc}^{\alpha,\alpha}|ai)^{(1)}$	-0.82	-0.43
$(ia|f_{xc}^{\alpha,\alpha}|ai)^{(2)}$	-0.83	-0.44

slight splitting of the $1^1\Pi$ excitation energy. This small effect is due to the fact that the grid used to calculate xc-integrals in DEMON2K has only roughly the symmetry of the molecule.

Now let us turn to BIGDFT calculations. Figure 9 shows a similar collapse of the continuum as the fineness of the grid increases. Interestingly there is no evidence of symmetry breaking of the doubly-degenerate $1^1\Pi$ state.

8.4. Oscillator Strengths

Carbon monoxide is very unusual for small molecules in that absolute oscillator strengths have been well studied [93] over a significant energy range and the $A^1\Pi$ ($1^1\Pi$ in Table 3) is bright and has an accurately determined oscillator strength. See Fig. 2 of Ref. [93] (as well as other references in the same paper) for a graph of measured absolute optical oscillator strengths against absorption energy in eV. Table 5 reports our calculated TD-LDA/TDA
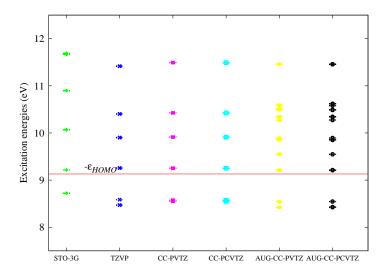
Figure 8. Singlet and triplet excitation energies for CO calculated using DEMON2K
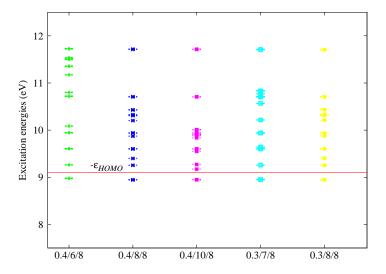

Figure 9. Singlet and triplet excitation energies for CO calculated using BIGDFT

oscillator strengths. As the TDA violates the Thomas-Reiche-Kuhn (TRK) f-sum rule [48] it should only be used very cautiously to estimate oscillator strengths. Nevertheless the DEMON2K value of $f = 0.232$ is in good agreement with the experimental value of $f = 0.1762$. As shown in Ref. [91], full TDLDA calculations with asymptotically corrected potentials give smaller oscillator strengths (0.136 for TD-LDA/LB94 and 0.156 for TD-LDA/AC-LDA calculations[91].) (Coïncidently our own DEMON2K full TD-LDA calcuations without asymptotic corrections give a degeneracyc-weighted oscillator strength of 0.1752 (bang on the experimental value) but an excitation energy of 8.19 eV.) Since os-

Table 5. Comparison of experimental $A^1\Pi$ energies (eV) and oscillator strengths with TD-LDA/TDA experimental $A^1\Pi$ energies (eV) and degeneracy-weighted oscillator strengths (unitless.)

	DEMON2K	BIGDFT	Experiment[a]
$\hbar\omega_S$	8.43	8.95	8.4
f	0.232	0.853	0.1762

[a] See Table VIII of Ref. [91].

cillator strengths are quite sensitive to configuration mixing with nearby states, the fact that the BIGDFT oscillator strength is larger than the DEMON2K oscillator strength may be due to the relatively small energy separation between the BIGDFT $A^1\Pi$ state and the artificially-low TD-LDA ionization continuum.

9.   Conclusion

In this chapter we have tried to give an informative elementary review of a subject largely unfamiliar to most theoretical chemists and physicists. Wavelets, once an obscure ripple at the exterior of engineering applications, grew to become a regular tsunami in engineering circles in the 1990s as the similarity to and superiority over Fourier transform methods for multiresolution problems with arbitrary boundary conditions became increasingly recognized. Though the first applications of wavelet theory to solving the Schrödinger equation may be traced back to the mid-1990s [94, 95, 96], the theory is still not well known among quantum mechanicians. Here we have tried to remedy this aberrant situation by trying to "make some *waves* about *wavelets* for *wave* functions."

In particular we have reviewed the theory behind the wavelet code BIGDFT for ground-state DFT and our recent implementation of wavelet-based TD-DFT in BIGDFT. Rapid progress is being made towards making BIGDFT a high performance computing order-N code for applications to large systems. Right now applications to 400 or 500 atoms are routine for ground-state calculations with BIGDFT. Our implementation of TD-DFT in BIGDFT is by comparison only a rudementary beginning, but it shows that the basic method is viable and we are confident that there are no insurmountable obstacles to making high performance computing order-N wavelet-based TD-DFT code for large systems.

Acknowledgments

Appendices

A List of Abbreviations

For the readers convenience we give a list of the abbreviations used in this chapter in alphabetical order:

AA Adiabatic approximation.

APW Augmented plane wave.

CC Coupled cluster.

CI Configuration interaction.

DFT Density-functional theory.

DM Density matrix.

FD Finite difference.

FE Finite element.

FEM Finite element method.

GTH-HGH Goedecker-Teter-Hutter/Hartwigsen-Goedecker-Hutter.

GTO Gaussian-type orbitals.

H Hartree.

HF Hartree-Fock.

HOMO Highest-occupied molecular orbital.

ISF Interpolating scaling function.

LCAO Linear combination of atomic orbitals.

LDA Local density approximation.

LMTO Linear muffin tin orbital.

LR Linear-response.

LR-TD-DFT Linear-response time-dependent density-functional theory.

LR-TD-LDA Linear-response time-dependent local density approximation.

LUMO Lowest-unoccupied molecular orbital.

NS Non-standard.

MBPT Many-body perturbation theory.

MRA Multiresolution analysis.

S Singlet.

SOS Sum-over-states.

STO Slater-type orbital

T Triplet.

TD Time-dependent.

TDA Tamm-Dancoff approximation.

TD-DFT Time-dependent density-functional theory.

TD-LDA Time-dependent local density approximation.

TRK Thomas-Reiche-Kuhn.

xc Exchange-correlation.

References

[1] T. H. Dunning, Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen, J. Chem. Phys. **90**, 1007 (1989).

[2] R. A. Kendall, T. H. Dunning, and R. J. Harrison, Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions, J. Chem. Phys. **96**, 6796 (1992).

[3] D. E. Woon and T. H. Dunning, Gaussian basis sets for use in correlated molecular calculations. III. The atoms aluminum through argon, J. Chem. Phys. **98**, 1358 (1993).

[4] X. P. Li, R. W. Nunes, and D. Vanderbilt, Density-matrix electronic-structure method with linear system-size scaling, Phys. Rev. B **47**, 10891 (1993).

[5] M. S. Daw, Model for energetics of solids based on the density matrix, Phys. Rev. B **47**, 10895 (1993).

[6] S. Goedecker and L. Colombo, Efficient linear scaling algorithm for tight-binding molecular dynamics, Phys. Rev. Lett. **73**, 122 (1994).

[7] W. Kohn, Density functional and density matrix method scaling linearly with the number of atoms, Phys. Rev. Lett. **76**, 3168 (1996).

[8] J. M. Millam and G. E. Scuseria, Linear scaling conjugate gradient density matrix search as an alternative to diagonalization for first principles electronic structure calculations, J. Chem. Phys. **106**, 5569 (1997).

[9] C. Ochsenfeld and M. Head-Gordon, A reformulation of the coupled perturbed self-consistent field equations entirely within a local atomic orbital density matrix-based scheme, Chem. Phys. Lett. **270**, 399 (1997).

[10] B. Huron, J. P. Malrieu, and P. Rancurel, Iterative perturbation calculations of ground and excited state energies from multiconfigurational zeroth-order wavefunctions, J. Chem. Phys. **58**, 5745 (1973).

[11] S. Evangelisti, J. P. Daudey, and J. P. Malrieu, Convergence of an improved CIPSI algorithm, Chem. Phys. **75**, 91 (1983).

[12] M. Schütz, G. Hetzer, and H. J. Werner, Low-order scaling local electron correlation methods. I. Linear scaling local MP2, J. Chem. Phys. **111**, 5691 (1999).

[13] G. E. Scuseria and P. Y. Ayala, Linear scaling coupled cluster and perturbation theories in the atomic orbital basis, J. Chem. Phys. **111**, 8330 (1999).

[14] D. R. Hartree, *The Calculation of Atomic Structures*, Wiley, New York, 1957.

[15] V. Fock, Näherungsmethode zur Lösung des quantenmechanischen Mehrkörperproblems, Z. für physik **61**, 126 (1930).

[16] J. C. Slater, Note on Hartree's method, Phys. Rev. **35**, 210 (1930).

[17] C. C. J. Roothan, New developments in molecular orbital theory, Rev. Mod. Phys. **23**, 69 (1951).

[18] L. H. Thomas, The calculation of atomic fields, Proc. Cambridge Phil. Soc. **23**, 542 (1926).

[19] E. Fermi, Sulla quantizzazione del gas perfetto monoatomico, Rend. Lincei **3**, 145 (1926).

[20] E. Fermi, Zur Quantelung des idealen einatomigen Gases, Z. für physik **36**, 902 (1926).

[21] J. Frenkel, Zur wellenmechanischen Theorie der metallischen Leitfähigkeit, Z. für physik **47**, 819 (1928).

[22] A. Sommerfeld, Zur Elektronentheorie der Metalle auf Grund der Fermischen Statistik I. Teil: Allgemeines, Strömungs- und Austrittsvorgänge, Z. für physik **47**, 1 (1928).

[23] P. A. M. Dirac, On the theory of quantum mechanics, Proc. Roy. Soc. London A **112**, 661 (1926).

[24] E. Fermi, Un metodo statistico per la determinazione di alcune priopretà dell'atomo, Rend. Accad. Naz. Lincei **6**, 602 (1927).

[25] P. A. M. Dirac, Note on exchange phenomena in the thomas atom, Math. Proc. Cambr. Phil. Soc. **26**, 376 (1930).

[26] P. Hohenberg and W. Kohn, Inhomogeneous electron gas, Phys. Rev. **136**, B864 (1964).

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

[28] R. M. Dreizler and E. K. U. Gross, *Density Functional Theory*, Springer-Verlag, Berlin Heidelberg New York, 1990.

[29] F. Herman, J. P. V. Dyke, and I. B. Ortenburger, Improved statistical exchange approximation for inhomogeneous many-electron systems, Phys. Rev. Lett. **22**, 807 (1969).

[30] O. K. Andersen, Linear methods in band theory, Phys. Rev. B **12**, 3060 (1975).

[31] E. Wimmer, H. Krakauer, M. Weinert, and A. J. Freeman, Full-potential self-consistent linearized-augmented-plane-wave method for calculating the electronic structure of molecules and surfaces: O_2 molecule, Phys. Rev. B **24**, 864 (1981).

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

[33] J. Pipek and I. Varga, Statistical electron densities, Int. J. Quant. Chem. **64**, 85 (1997).

[34] R. Poirier, R. Kari, and I. G. Csizmadia, *Handbook of Gaussian Basis sets*, Elsevier, Amsterdam, 1970.

[35] J. A. Pople and D. L. Beveridge, *Approximate Molecular Orbital Theory*, McGraw-Hill, New York, 1970.

[36] S. Wei and M. Y. Chou, Wavelets in self-consistent electronic structure calculations, Phys. Rev. Lett. **76**, 2650 (1996).

[37] T. Kato, On the eigenfunctions of many particle systems in quantum mechanics, Comm. Pure Appl. Math **10**, 151 (1957).

[38] T. A. Arias, R. A. Lippert, and A. Edelman, Multiscale computation with interpolating wavelets, J. Comput. Phys. **140**, 278 (1997).

[39] J. C. Slater, *Quantum Theory of Atomic Structure*, McGraw-Hill, New York, 1960.

[40] C. F. Fischer, Numerical solution of the Hartree-Fock equations, Can. J. Phys. **41**, 1895 (1963).

[41] C. F. Fischer, Average-energy-of-configuration Hartree-Fock results for the atoms helium to radon charlotte froese fischer, At. Data Nucl. Data Tables **4**, 301 (1972).

[42] C. F. Fischer, Average-energy-of-configuration Hartree-Fock results for the atoms helium to radon, At. Data Nucl. Data Tables **12**, 87 (1973).

[43] J. B. Mann, SCF Hartree-Fock results for elements with two open shells and the elements francium to nobelium, At. Data Nucl. Data Tables **12**, 1 (1973).

[44] D. Heinemann, B. Ficke, and D. Kolb, Accurate Hartree-Fock-Slater calculations on small diatomic molecules with the finite-element method, Chem. Phys. Lett. **145**, 125 (1988).

[45] B. Natarajan, L. Genovese, M. E. Casida, T. Deutsch, O. N. Burchak, C. Philouze, and M. Y. Balakirev, Wavelet-based linear-response time-dependent density-functional theory, http://arxiv.org/abs/1108.3475.

[46] E. Runge and E. K. U. Gross, Density-functional theory for time-dependent systems, Phys. Rev. Lett. **52**, 997 (1984).

[47] M. Petersilka, U. J. Gossmann, and E. K. U. Gross, Excitation energies from time-dependent density-functional theory, Phys. Rev. Lett. **76**, 1212 (1996).

[48] M. E. Casida, Time-dependent density-functional response theory for molecules, in *Recent Advances in Density Functional Methods, Part I*, edited by D. P. Chong, page 155, World Scientific, Singapore, 1995.

[49] C. Jamorski, M. E. Casida, and D. R. Salahub, Dynamic polarizabilities and excitation spectra from a molecular implementation of time-dependent density-functional response theory: N_2 as a case study, J. Chem. Phys. **104**, 5134 (1996).

[50] R. Bauernschmitt and R. Ahlrichs, Treatment of electronic excitations within the adiabatic approximation of time-dependent density functional theory, Chem. Phys. Lett., **256**, 454 (1996).

[51] C. Bienia, S. Kumar, J. P. Singh, and K. Li, The PARSEC benchmark suite: Characterization and architectural implications, Technical report, in Princeton University, 2008.

[52] A. Castro, M. A. L. Marques, H. Appel, M. Oliveira, C. A. Rozzi, X. Andrade, F. Lorenzen, E. K. U. Gross, and A. Rubio, Octopus: A tool for the application of time-dependent density functional theory, Physica Status Solidi **243**, 2465 (2006).

[53] L. Lehtovaara, V. Havu, and M. Puska, All-electron density functitonal theory and time-dependent density-functional theory with high-order finite elements, J. Chem. Phys. **131**, 054103 (2009).

[54] O. M. Nielsen, *Wavelets in scientific computing*, PhD thesis, Department of Mathematical Modelling, Technical University of Denmark, 1998.

[55] W. Dahmen, A. Cohen, and R. DeVore, Adaptive wavelet schemes for elliptic operator equations - convergence rates, Math. Comput. **70**, 27 (2001).

[56] T. K. Jensen, *On adaptive wavelet-based methods for the Maxwell equations*, PhD thesis, Department of Mathematics, Technical University of Denmark, 2003.

[57] http://inac.cea.fr/L_Sim/BigDFT/.

[58] F. Keinert, *Wavelets and Multiwavelets*, Chapman and Hall/CRC, 2003.

[59] A. Haar, Zur Theorie des orthogonalen Funktionensysteme, Mathematische Annalen **69**, 331 (1910).

[60] B. K. Alpert, A class of bases in L^2 for the sparse representation of integral operators, SIAM J. Math. Anal. **24**, 246 (1993).

[61] I. Daubechies, *Ten Lectures on Wavelets*, SIAM, Philadelphia, 1992.

[62] R. A. Lippert, T. A. Arias, and A. Edelman, Multiscale computation with interpolating wavelets, J. Comput. Phys. **140**, 278 (1998).

[63] P. Pulay, Convergence acceleration of iterative sequences. The case of SCF iteration, Chem. Phys. Lett. **73**, 393 (1980).

[64] E. R. Davidson, The iterative calculation of a few of the lowest eigenvalues and corresponding eigenvectors of large real-symmetric matrices, J. Comput. Phys. **17**, 87 (1975).

[65] C. W. Murray, S. C. Racine, and E. R. Davidson, Improved algorithms for the lowest few eigenvalues and eigenvectors of large matrices, J. Comput. Phys. **103**, 382 (1991).

[66] B. Walker, A. M. Saitta, R. Gebauer, and S. Baroni, Efficient approach to time-dependent density-functional perturbation theory for optical spectroscopy, Phys. Rev. Lett. **96**, 113001 (2006).

[67] D. Rocca, R. Gebauer, Y. Saad, and S. Baroni, Turbo charging time-dependent density-functional theory with Lanczos chains, J. Chem. Phys. **128**, 154105 (2008).

[68] S. Baroni, R. Gebauer, O. B. Malcioğlu, Y. Saad, P. Umari, and J. Xian, Harnessing molecular excited states with Lanczos chains, J. Phys. Condens. Matter **22**, 074204 (2010).

[69] O. B. Malcioğlu, R. Gebauer, D. Rocca, and S. Baroni, TURBOTDDFT − A code for the simulation of molecular spectra using the Liouville-Lanczos approach to time-dependent density-functional perturbation theory, Comp. Phys. Comm. **182**, 1744 (2011).

[70] DEMON2K @ GRENOBLE, the Grenoble development version of DEMON2K, Andreas M. Köster, Patrizia Calaminici, Mark E. Casida, Roberto Flores-Morino, Gerald Geudtner, Annick Goursot, Thomas Heine, Andrei Ipatov, Florian Janetzko, Sergei

Patchkovskii, J. Ulisis Reveles, Dennis R. Salahub, and Alberto Vela, *The International deMon Developers Community* (Cinvestav-IPN, Mexico, 2006) *plus some additional features* by Mark E. Casida, Loïc Joubert Doriol, Andrei Ipatov, Miquel Huix-Rotllant, and Bhaarathi Natarajan (Grenoble, France, 2011).

[71] L. Genovese, T. Deutsch, A. Neelov, S. Goedecker, and G. Beylkin, Efficient solution of poisson's equation with free boundary conditions, J. Chem. Phys., **125**, 074105 (2006).

[72] L. Genovese, T. Deutsch, and S. Goedecker, Efficient and accurate three-dimensional poisson solver for surface problems, J. Chem. Phys. **127**, 054704 (2007).

[73] L. Genovese et al., Daubechies wavelets as a basis set for density functional pseudopotential calculations, J. Chem. Phys. **129**, 014109 (2008).

[74] L. Genovese, M. Ospici, T. Deutsch, J.-F. Méhaut, A. Neelov, and S. Goedecker, Density functional theory calculation on many-cores hybrid cpu-gpu architectures, J. Chem. Phys. **131**, 034103 (2009).

[75] T. Deutsch and L. Genovese, Wavelets for electronic structure calculations, Journées des Neutrons **18**, 33 (2011).

[76] L. Genovese, B. Videau, M. Ospici, T. Deutsch, S. Goedecker, and J.-F. Méhaut, Daubechies wavelets for high performance electronic structure calculations: The bigdft project, Comptes Rendus Mécanique **339**, 149 (2011).

[77] G. Beylkin, On the representation of operators in bases of compactly supported wavelets, SIAM J. Numer. Anal. **6**, 1716 (1992).

[78] C. J. Tymczak and X.-Q. Wang, Orthonormal wavelet bases for quantum molecular dynamics, Phys. Rev. Lett. **78**, 3654 (1997).

[79] A. I. Neelov and S. Goedecker, An efficient numerical quadrature for the calculation of the potential energy of wavefunctions expressed in the Daubechies wavelet basis, J. Comp. Phys. **217**, 055501 (2006).

[80] K.-A. Lau and Q. Sun, Asymptotic regularity of Daubechies scaling functions, Proc. Am. Math. Soc. **128**, 1087 (2000).

[81] S. Goedecker, *Wavelets and Their Application for the Solution of Partial Differential Equations*, Presses Polytechniques Universitaires et Romandes, Lausanne, Switzerland, 1998.

[82] S. Goedecker, M. Teter, and J. Hutter, Separable dual-space Gaussian pseudopotentials, Phys. Rev. B **54**, 1703 (1996).

[83] S. Goedecker, M. Teter, and J. Hutter, Relativistic separable dual-space Gaussian pseudopotentials from H to Rn, Phys. Rev. B **58**, 3641 (1998).

[84] B. R. Johnson, J. P. Modisette, P. J. Nordlander, and J. L. Kinsey, Quadrature integration for orthogonal wavelet systems, J. Chem. Phys. **110**, 8309 (1999).

[85] D. Feller, The role of databases in support of computational chemistry calcuations, J. Comp. Phys. **17**, 1571 (1996).

[86] K. L. Schuchardt, B. T. Didier, T. Elsethagen, L. Sun, V. Gurumoorthi, J. Chase, J. Li, and T. L. Windus, Basis set exchange: A community database for computational sciences, J. Chem. Inf. Model. **47**, 1045 (2007).

[87] A. Ipatov, A. Fouqueau, C. P. del Valle, F. Cordova, M. E. Casida, A. M. Köster, A. Vela, and C. J. Jamorski, Excitation energies from an auxiliary-function formulation of time-dependent density-functional response theory with charge conservation constraint, J. Molec. Struct. (Theochem) **762**, 179 (2006).

[88] M. E. Casida, F. Gutierrez, J. Guan, F. Gadea, D. R. Salahub, and J. Daudey, Charge-transfer correction for improved time-dependent local density approximation excited-state potential energy curves: Analysis within the two-level model with illustration for H_2 and LiH, J. Chem. Phys. **113**, 7062 (2000).

[89] M. E. Casida, Review: Time-dependent density-functional theory for molecules and molecular solids, J. Mol. Struct. (Theochem) **914**, 3 (2009).

[90] M. E. Casida and M. Huix-Rotllant, Progress in time-dependent density-functional theory, Annu. Rev. Phys. Chem. **63**, in press.

[91] M. E. Casida and D. R. Salahub, Asymptotic correction approach to improving approximate exchange-correlation potentials: Time-dependent density-functional theory calculations of molecular excitation spectra, J. Chem. Phys., **113**, 8918 (2000).

[92] M. E. Casida, C. Jamorski, K. C. Casida, and D. R. Salahub, Molecular excitation energies to high-lying bound states from time-dependent density-functional response theory: Characterization and correction of the time-dependent local density approximation ionization threshold, J. Chem. Phys. **108**, 4439 (1998).

[93] W. F. Chan, G. Cooper, and C. E. Brion, Absolute optical oscillator strengths for discrete and continuum photoabsorption of carbon monoxide (7-200 ev) and transition moments for the $X^1\Sigma^+ \to A^1\Pi$ system, Chem. Phys. **170**, 123 (1993).

[94] T. A. Arias, Multiresolution analysis of electronic structure: semicardinal and wavelet bases, Rev. Mod. Phys. **71**, 267 (1999).

[95] P. Fischer and M. Defranceschi, Looking at atomic orbitals through Fourier and wavelet transforms, Int. J. Quant. Chem. **45**, 619 (1993).

[96] J.-L. Calais, Wavelets–Something for quantum chemistry?, Int. J. Quant. Chem. **58**, 541 (1996).