A new future for finite-element methods in Quantum Chemistry?

Kenneth Ruud

Department of Chemistry
University of Tromsø
9037 Tromsø
Norway

Geilo, 29/1 2007

Historical viewpoints on mathematics in chemistry

> “Every attempt to employ mathematical methods in the study of chemical questions must be considered profoundly irrational. If mathematical analysis should ever hold a prominent place in chemistry—an aberration which is happily impossible—it would occasion a rapid and widespread degradation of that science.”
> *Auguste Comte, Cours de philosophie positive, 1830*

Historical viewpoints on mathematics in chemistry

“Every attempt to employ mathematical methods in the study of chemical questions must be considered profoundly irrational. If mathematical analysis should ever hold a prominent place in chemistry—an aberration which is happily impossible—it would occasion a rapid and widespread degradation of that science.”
Auguste Comte, Cours de philosophie positive, 1830

“The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble.”
Paul Dirac, Quantum mechanics of many electron systems, 1929

Historical viewpoints on mathematics in chemistry

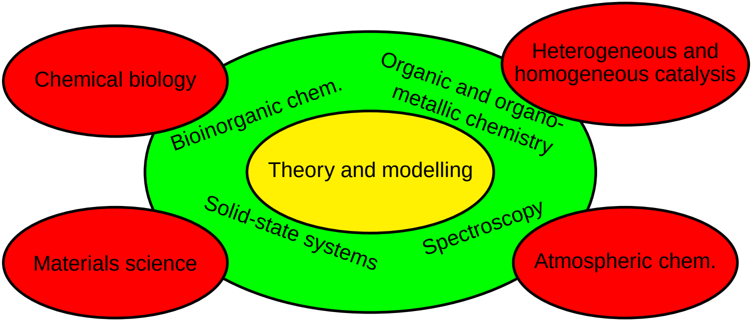
"Every attempt to employ mathematical methods in the study of chemical questions must be considered profoundly irrational. If mathematical analysis should ever hold a prominent place in chemistry—an aberration which is happily impossible—it would occasion a rapid and widespread degradation of that science."
Auguste Comte, Cours de philosophie positive, 1830

"The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble."
Paul Dirac, Quantum mechanics of many electron systems, 1929

"Quantum chemistry is today used within all branches of chemistry and molecular physics. As well as producing quantitative information on molecules and their interactions, the theory also affords deeper understanding of molecular processes that cannot be obtained from experiments alone."
The Royal Swedish Academy of Sciences, Press release: The 1998 Nobel Prize in Chemistry

Our vision for the future

- The Center of Theoretical and Computational Chemistry



Chemical biology

Heterogeneous and homogeneous catalysis

Organic and organo-metallic chemistry

Bioinorganic chem.

Theory and modelling

Solid-state systems

Spectroscopy

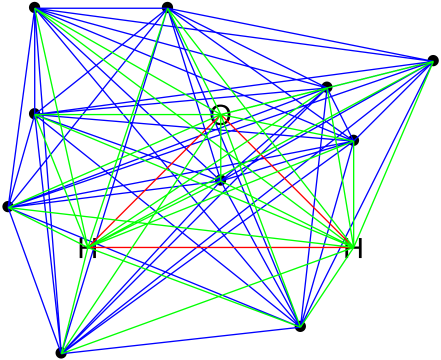
Materials science

Atmospheric chem.

- Evenly split between the Universities of Oslo and Tromsø
- 10 theoretical and computational scientists including 2 YFFs and Norways most highly cited chemist

The quantum-mechanical view of a molecule

- An instantaneous pictures of the interactions present in the water molecule (3 nuclei and 10 electrons)



The electronic Schrödinger equation

- The electrons in the molecule are described by a wave function Ψ determined by solving the electronic Schrödinger equation:

$$H \quad \Psi \quad = E \quad \Psi$$

The electronic Schrödinger equation

- The electrons in the molecule are described by a wave function Ψ determined by solving the electronic Schrödinger equation:

$$\enclose{circle}{H}\ \Psi = E\ \Psi$$

- H is the electronic Hamiltonian operator

$$H = -\frac{e^2}{2m_e}\sum_i \nabla_i^2 - e^2 \sum_{iK} \frac{Z_K}{R_{iK}} + \frac{e^2}{2}\sum_{ij}\frac{1}{r_{ij}} + V_{\mathrm{NUC}}$$

The electronic Schrödinger equation

- The electrons in the molecule are described by a wave function Ψ determined by solving the electronic Schrödinger equation:

$$H \enclose{circle}{\Psi} = E \enclose{circle}{\Psi}$$

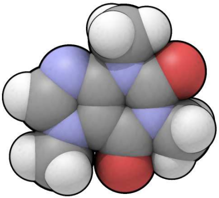
- H is the electronic Hamiltonian operator

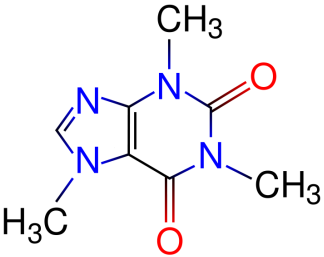
$$H = -\frac{e^2}{2m_e}\sum_i \nabla_i^2 - e^2 \sum_{iK} \frac{Z_K}{R_{iK}} + \frac{e^2}{2}\sum_{ij}\frac{1}{r_{ij}} + V_{\mathrm{NUC}}$$

- The wave function Ψ is a function of $3N$ variables (where N is the number of electrons in the molecule)

The quantum-mechanical challenge 1

- The Schrödinger equation is a complicated many-particle equation
- The only atom we can solve exactly (non-relativistic) is the hydrogen atom (one electron and one nucleus)
- Even a small organic molecule such as caffeine contains 136 particles (24 nuclei and 102 electrons)





- All particles interact with each other, and we can only determine approximative solutions to the Schrödinger equation

The quantum-mechanical challenge 2

- Chemists are demanding customers

The quantum-mechanical challenge 2

- Chemists are demanding customers
- The total molecular energy of caffeine is 1 800 000 kcal/mol.
- The energy of the chemical bond (which we would like to determine with an accuracy of 1%) is 300-400 kcal/mol—that is, the bond energy is 0.02% of the total energy

The quantum-mechanical challenge 2

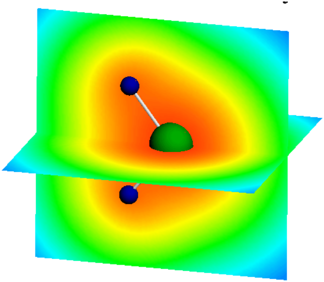
- Chemists are demanding customers
- The total molecular energy of caffeine is 1 800 000 kcal/mol.
- The energy of the chemical bond (which we would like to determine with an accuracy of 1%) is 300-400 kcal/mol—that is, the bond energy is 0.02% of the total energy
- Quantum-mechanically we thus have to determine the weight of a boat captain by weighing the boat with and without the captain

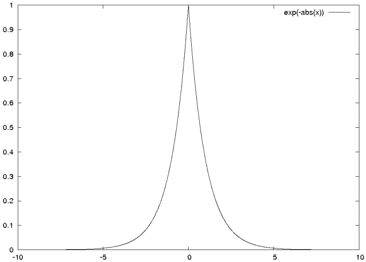
Ye, god's! You're right, Captain! We are in a bottle!

The electron density

- The wave function itself is not observable, but the electron density of the molecule is:

$$\rho\left(\mathbf{r}\right) = \int_{\nu_2}\int_{\nu_3}\dots\int_{\nu_N} \Psi\left(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N\right)^* \Psi\left(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N\right) d\mathbf{r}_2 d\mathbf{r}_3 \dots \mathbf{r}_N$$





- Ψ thus contains too much information whereas $\rho\left(\mathbf{r}\right)$ is a function of only three variables

Kohn-Sham formulation of DFT

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

$$\left[-\frac{1}{2}\nabla^2 + V_{nuc} + V_{coul}\left(\rho\right) + V_{xc}\left(\rho\right)\right]\phi_i(r) = \epsilon_i\phi_i(r) \tag{2}$$

$$V_{nuc} = \sum_{i=1}^{N} Z_i / \left|r - R_i\right| \tag{3}$$

$$V_{coul} = \int \mathrm{d}r' \frac{\rho(r')}{|r - r'|} \tag{4}$$

$$V_{xc} = f(\rho(r)) \tag{5}$$

The Traditional Quantum Chemical Approach

- Express the MOs as a linear combination of AOs located on the nuclei of the molecule

$$\phi_i(r) = \sum_j C_{ij}\chi_j \qquad \chi_j = x_A^m y_A^n z_A^o \exp\left(-\alpha\left(x_A^2 + y_A^2 + z_A^2\right)\right) \tag{6}$$

- Represent the operators and density in terms of the AOs
- Rewrite the Kohn-Sham equations as an SCF problem:

$$F(C)\,C = SC\epsilon \tag{7}$$

Possible problems:

- Scaling with the system size (at least N^2 or N^3)
- Non-orthogonality of the basis-set
- Expensive evaluation of integrals

Why argue with success?

- Quantum-chemical methods do not show the expected linear scaling
- In order to achieve linear scaling, the methods should ideally be inherently sparse and banded
- Delocalized molecular orbitals and symmetry-adapted atomic orbitals not optimal
- It would also be convenient to be able to easily do divide-and-conquer methods
- Potential problems in supersystem calculations are:
  - Arbitrary boundaries between QM and MM systems
  - BSSE errors between QM systems
  - Important interaction mechanisms can be lost with different computational methods between interaction systems
- A possible new pathway: multiwavelet bases?

Integral formulation of the K-S equations

- We can rewrite the KS equations in an integral formulation

$$\phi_i(r) = -2 \int \mathrm{d}r' G_\mu(r, r') V(r') \phi_i(r') \tag{8}$$

where

$$G_\mu(r, r') = \frac{1}{4\pi} \frac{e^{-\mu|r-r'|}}{|r-r'|} \qquad \mu = \sqrt{-2\epsilon_i} \tag{9}$$

- The standard K-S equations are recovered noticing that

$$(-\nabla_r^2 + \mu^2) G_\mu(r, r') = \delta(r, r') \tag{10}$$

- The integral formulation important to avoid divergencies due to ∇^2 in finite-element methods

Scaling functions: 1-D basis in [0,1]

Scaling functions: Legendre polynomials of order k in [0,1]

$$\phi_k(x) = (2k+1)^{1/2} P_k(2x-1) \qquad 0 \leq x \leq 1 \tag{11}$$

Translation and dilation

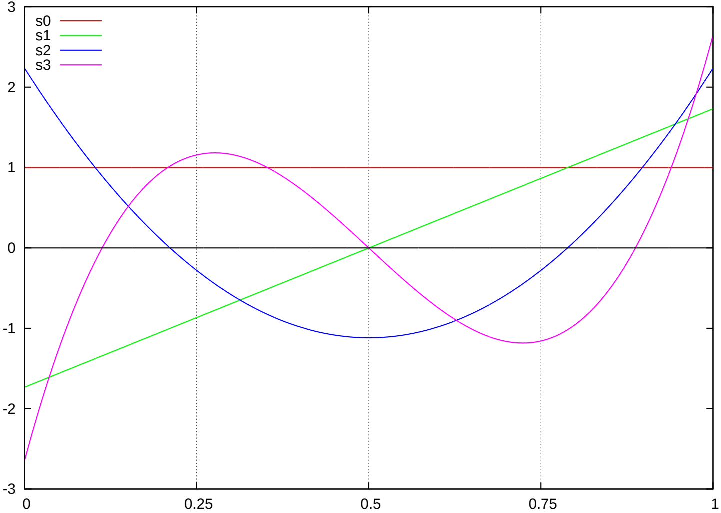
$$\phi^n_{k,l}(x) = 2^{n/2}\phi_k(2^n x - l) \qquad 2^{-n}l \leq x \leq 2^{-n}(l+1) \tag{12}$$

Orthonormality (Legendre polynomials + disjoint support)

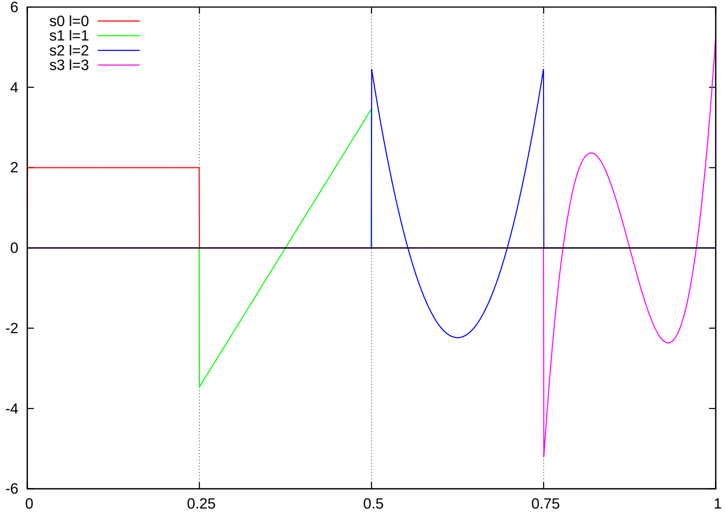
$$\int_0^1 \mathrm{d}x\, \phi^n_{k,l}(x)\phi^n_{k',l'}(x) = \delta_{kk'}\delta_{ll'} \tag{13}$$

“Ladder” of spaces: $V_0 \subset V_1 \subset V_2 \subset \ldots V_n \subset \ldots$
This basis is complete in $L_2[0,1]$ both in the limit of infinite n and infinite k

Example: scaling functions up to order $k = 3$



Example: scaling functions up to order $k = 3$ at $n = 2$ and different translations



Projection of a function in 1-D

Given a function $f(x)$ defined in $[0,1]$, its projection on V_n can be written as:

$$\tilde{f}(x) = \sum_{l=0}^{2^n-1} \sum_{j=0}^{k} s_{jl}^n \phi_{j,l}^n(x) \tag{14}$$

$$s_{jl}^n = \int_{2^{-n}l}^{2^{-n}(l+1)} f(x)\phi_{j,l}^n(x) \tag{15}$$

Quadrature employed to evaluate the integral

$$s_{jl}^n = \sum_{i=0}^{k} w_i f(x_i)\phi_{i,l}^n(x_i) \tag{16}$$

The projection at level n is in general only approximate.
What is the error in the projection?

Wavelet functions

- Definition: W_k^n is the orthogonal complement of V_k^n in V_k^{n+1}

$$V^n \oplus W^n = V^{n+1} \tag{17}$$

- Recursively::

$$V^n = V_0 \oplus W^1 \oplus \ldots W^{n-1} \tag{18}$$

- The functions $[\psi_{k,l}^n(x)]$, constituting a basis for W_n, are called *wavelet functions*.

Properties of the wavelet functions

- Translation and dilation:

$$\psi_{i,l}^{n}(x) = 2^{n/2}\psi_i(2^n x - l)$$

- $k+1$ vanishing moments:

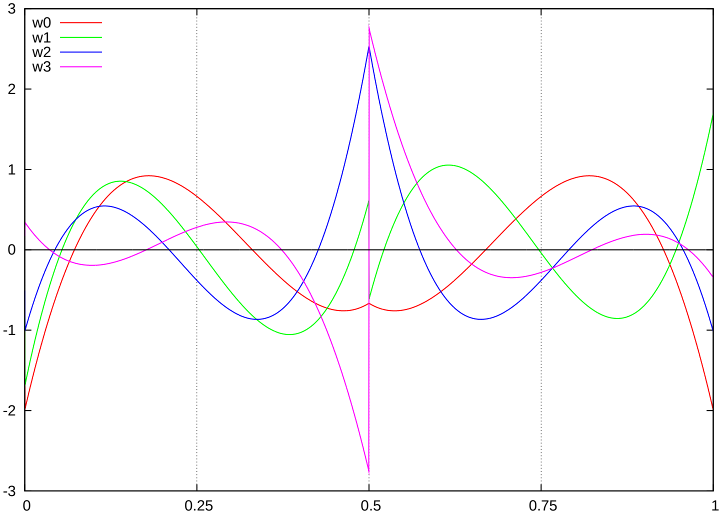
$$\int x^{\alpha}\psi_{i,l}^{n}(x)\,\mathrm{d}x = 0 \qquad \alpha = 0, 1 \ldots k$$

- Orthogonality:

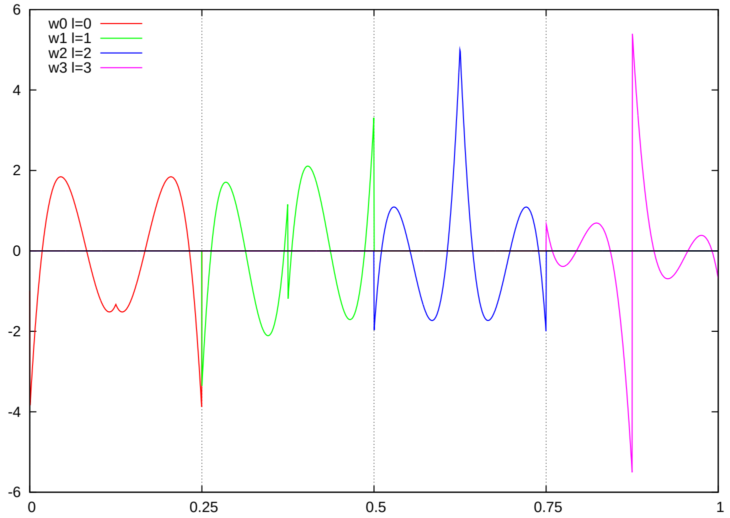
$$\int \psi_{i,k}^{n}(x)\psi_{j,m}^{m}(x)\,\mathrm{d}x = \delta_{ij}\delta_{kl}\delta_{nm}$$

- Choice of wavelets is not unique unless additional constraints are added

Example: wavelet functions up to order $k = 3$



Example: wavelet functions up to order $k = 3$ at $n = 2$ and different translation



Two-scale relations and filters

- Both scaling and wavelet functions at scale n can be constructed from the scaling functions at scale $n+1$:

$$\phi^n_{i,l}(x) = 2^{1/2} \sum_j h^0_{ij} \phi^{n+1}_{j,2l}(x) + h^1_{ij} \phi^{n+1}_{j,2l+1}(x) \tag{19}$$

$$\psi^n_{i,l}(x) = 2^{1/2} \sum_j g^0_{ij} \phi^{n+1}_{j,2l}(x) + g^1_{ij} \phi^{n+1}_{j,2l+1}(x) \tag{20}$$

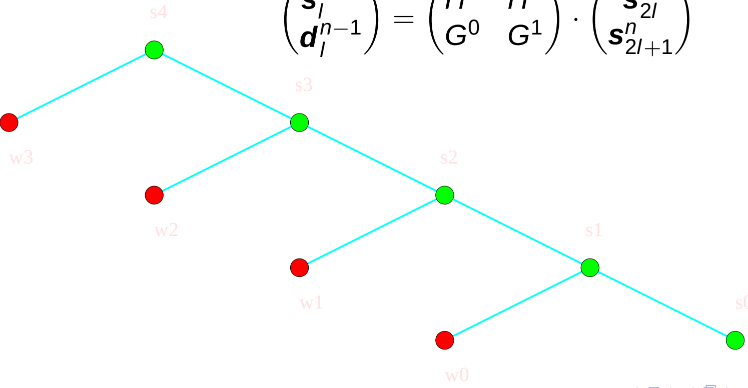
- More concisely, we can write this as a wavelet filter involving matrix-vector products

$$\begin{pmatrix} \phi^n_l(x) \\ \psi^n_l(x) \end{pmatrix} = 2^{1/2} \begin{pmatrix} H^0 & H^1 \\ G^0 & G^1 \end{pmatrix} \cdot \begin{pmatrix} \phi^{n+1}_{2l}(x) \\ \phi^{n+1}_{2l+1}(x) \end{pmatrix} \tag{21}$$

Multiwavelet decomposition

Once $f(x)$ has been projected on the finest scale n, the scaling and wavelet coefficients at coarser scales ($m < n$) are obtained by applying the filters:

$$\begin{pmatrix} \mathbf{s}_l^{n-1} \\ \mathbf{d}_l^{n-1} \end{pmatrix} = \begin{pmatrix} H^0 & H^1 \\ G^0 & G^1 \end{pmatrix} \cdot \begin{pmatrix} \mathbf{s}_{2l}^n \\ \mathbf{s}_{2l+1}^n \end{pmatrix} \tag{22}$$



Representations of a function in 1-D

Projected (or reconstructed) representation

$$\tilde{f}(x) = \sum_{l=0}^{2^n-1} \sum_{j=0}^{k} s_{jl}^{n} \phi_{j,l}^{n}(x) \tag{23}$$

Compressed representation obtained by recurring along the “ladder”:

$$\tilde{f}(x) = \sum_{j=0}^{k} \left(s_{j,0}^{0} \phi_j(x) + \sum_{m=0}^{n-1} \sum_{l=0}^{2^m-1} d_{jl}^{m} \psi_{j,l}^{n}(x) \right) \tag{24}$$

In practice both the scaling (s_{il}^n) and the wavelet (d_{il}^n) are stored.

Accuracy and adaptive refinement

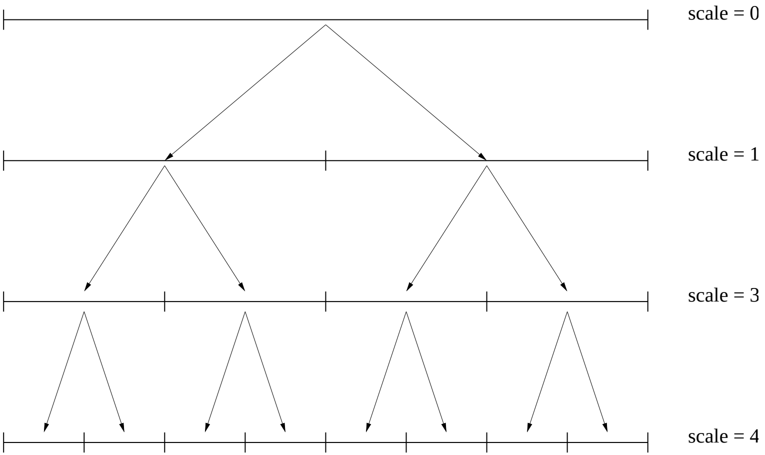
We require arbitrary and predefined accuracy:

$$\left|\tilde{f}(x) - f(x)\right| < \epsilon \quad \forall 0 \leq x \leq 1 \tag{25}$$

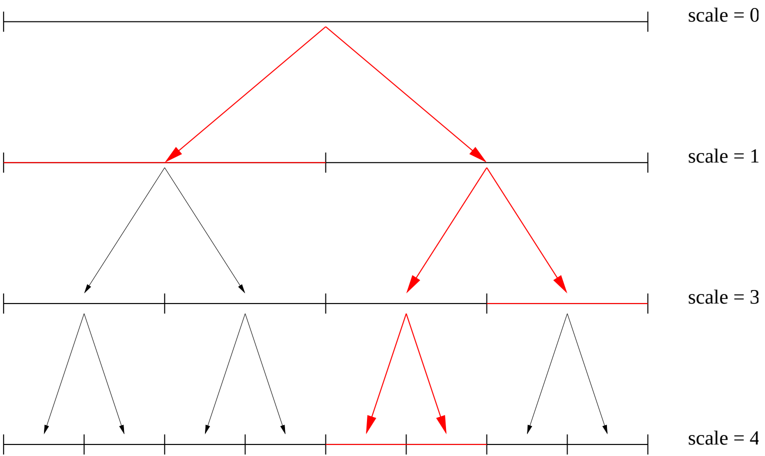
We need an algorithm to obtain the representation of a function
In practice the following algorithm is used:

1. For a given node n, l compute $s^n_{j,l}$ and $d^n_{j,l}$;
2. Check if $\|d^n_{j,l}\| \leq \epsilon$;
3. If fulfilled keep the node (no subdivision);
4. If not fulfilled, subdivide the node and go to n.1 for $(n+1, 2l)$ and $n+1, 2l+1$.

Uniform Refinement

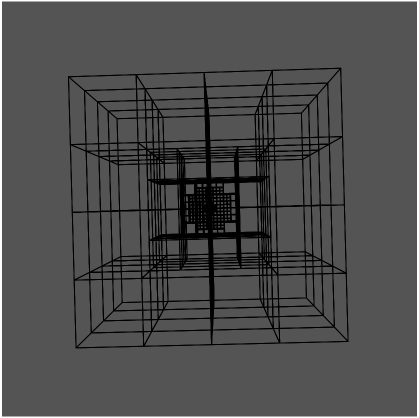


Adaptive Refinement



An example of an adaptive grid

The projection of $f(x_1, x_2, x_3) = c^{-1} e^{-500((x_1-0.5)^2+(x_2-0.5)^2+(x_3-0.5)^2)}$, up to relative precision $\|f - \hat{f}\|_2 / \|f\|_2 < \epsilon = 10^{-8}$



A total of 649 grid point needed

NS-form of an operator

Given an operator T such that

$$g(y) = T \cdot (f(x)) \tag{26}$$

Projection:

$$T^n = P^n T P^n \tag{27}$$

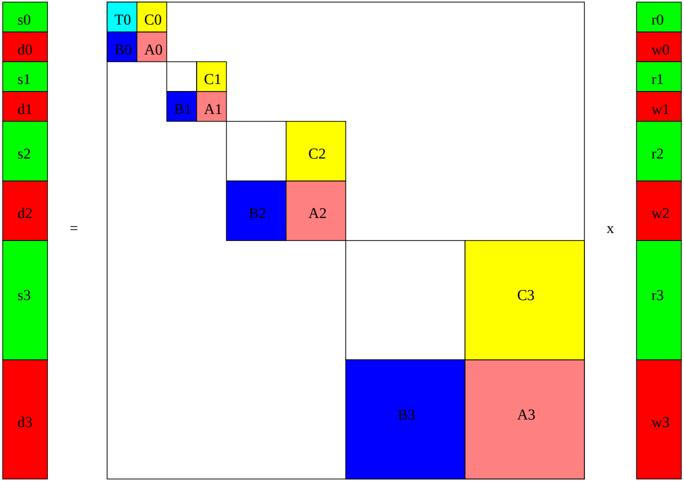
We can express the project using both scaling and wavelet functions

$$T^{n+1} = P^{n+1} T P^{n+1} = \left(P^n + Q^n\right) T \left(P^n + Q^n\right) \tag{28}$$

We can then write write a telescopic series

$$\begin{aligned} T^{n+1} &= T^0 + \sum_{m=1}^{n} \left(Q^m T Q^m + Q^m T P^m + P^m T Q^m\right) \\ &= T^0 + \sum_{m=1}^{n} \left(A^m + B^m + C^m\right) \end{aligned}$$

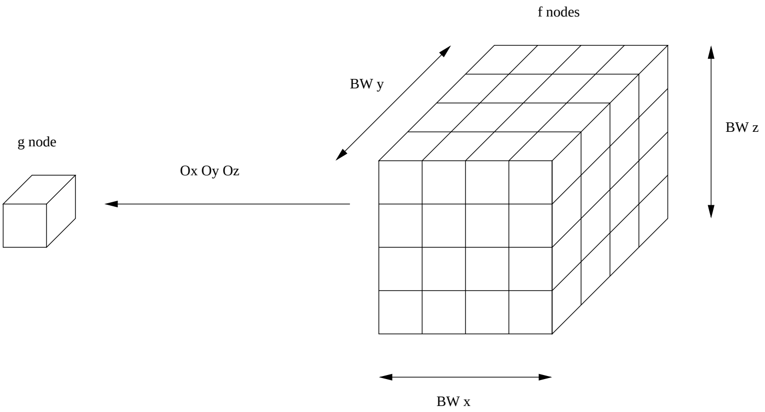
Application of an operator



A, *B*, and *C* are very narrow banded along the main diagonal (thresholding).

T is dense but needed only at the coarsest scale (one node only!).

3d application of the banded operator



On the cost of applying an operator

For a given polynomial order k:

- In 1-D the speed is $O(k^2)$ (fast!).
- In 3-D the speed is $O(k^6)$ (too expensive!) for a straightforward implementation
- Exploiting kernel separability $K(x,y,z) = K'(x)K'(y)K'(z)$ $O(k^4)$ (feasible).
- Problem: the kernels needed are not separable (formally) but ...

$$K_P(\mathbf{x}-\mathbf{y}) = \frac{1}{|\mathbf{x}-\mathbf{y}|} \qquad K_H(\mathbf{x}-\mathbf{y}) = \frac{e^{-\mu|\mathbf{x}-\mathbf{y}|}}{|\mathbf{x}-\mathbf{y}|} \tag{29}$$

Separated representation of the Poisson and Helmholtz kernels

The MW approach is feasible only if it is possible to write the kernel of an operator as

$$K(\boldsymbol{x}-\boldsymbol{y}) = \sum_{l=1}^{M} c_l K_l(x_1-y_1) K_l(x_2-y_2) K_l(x_3-y_3) \tag{30}$$

For DFT in its integral formulation, the kernels to be used are:

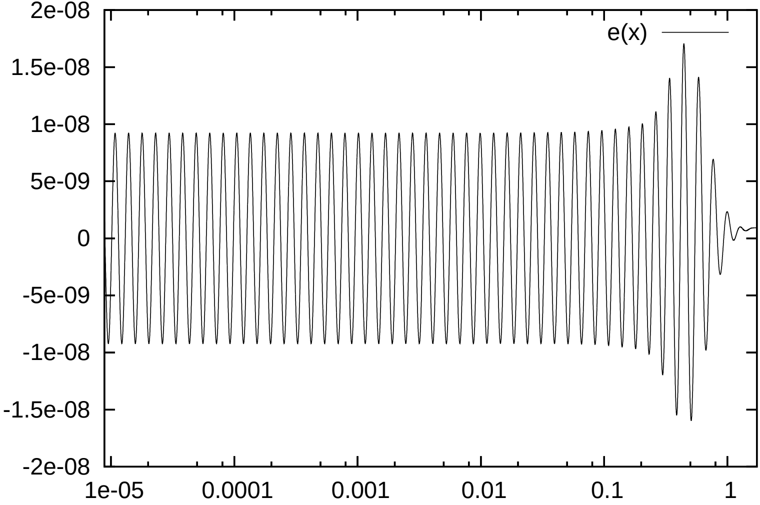
- The Poisson kernel: $K_P(\boldsymbol{x}-\boldsymbol{y}) = 1/\left|\boldsymbol{x}-\boldsymbol{y}\right|$
- The Helmholtz kernel: $K_H(\boldsymbol{x}-\boldsymbol{y}) = e^{-\mu|\boldsymbol{x}-\boldsymbol{y}|}/\left|\boldsymbol{x}-\boldsymbol{y}\right|$

For a predefined precision ϵ it is possible to obtain an expansion such that:

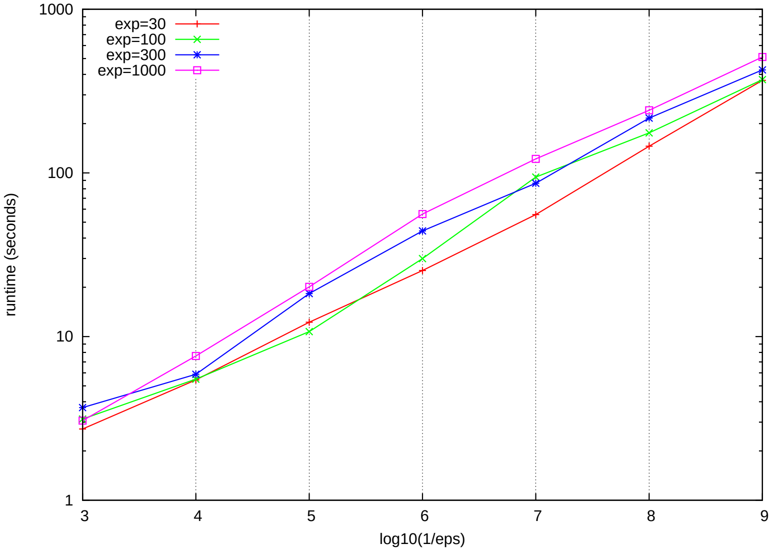
$$\left| K(\boldsymbol{x}-\boldsymbol{y}) - \sum_{l=1}^{M} c_l e^{-\omega_l(\boldsymbol{x}-\boldsymbol{y})^2} \right| \leq \epsilon \tag{31}$$

Separated Poisson kernel: relative error

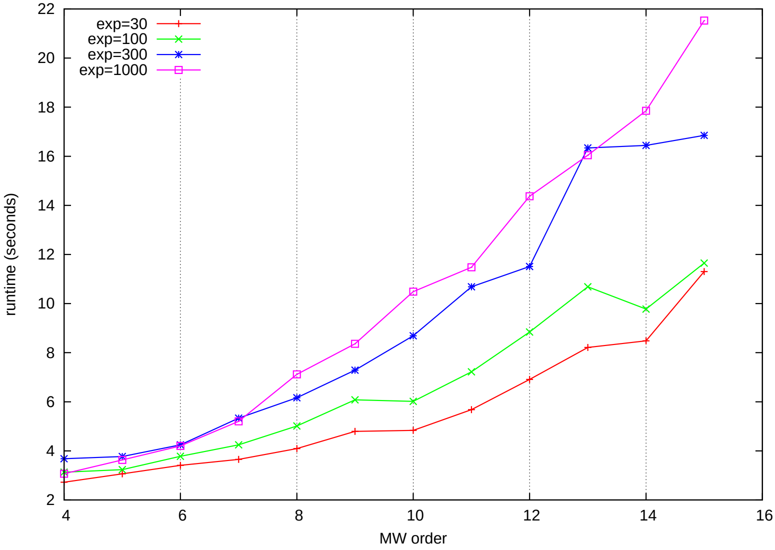
70 terms, accuracy 10^{-8}



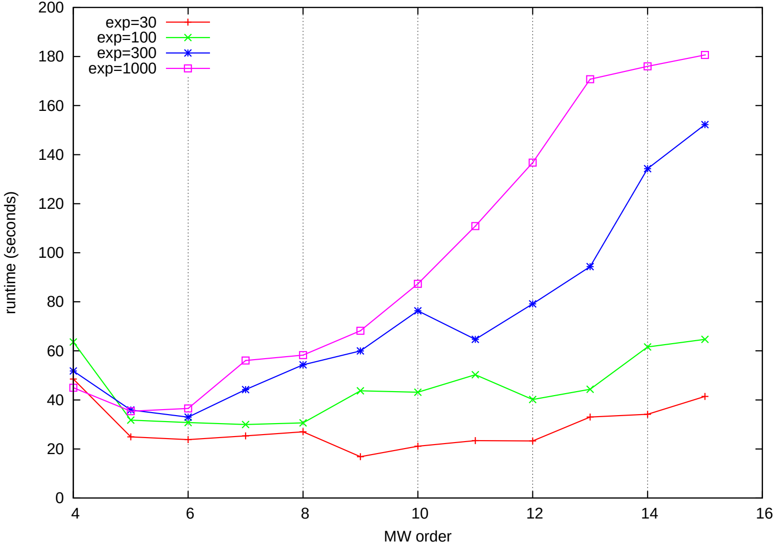
Poisson kernel for different accuracies



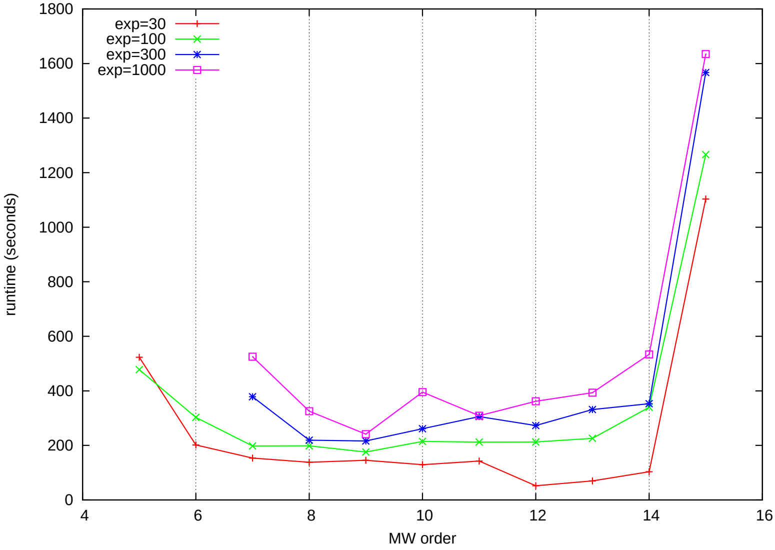
Poisson kernel for fixed resolution $\epsilon = 10^{-3}$



Poisson kernel for fixed resolution $\epsilon = 10^{-6}$



Poisson kernel for fixed resolution $\epsilon = 10^{-8}$



Summary

- A new approach with several interesting properties: scaling, orthonormality, predefined accuracy
- Inherently linear scaling, and probably quite parallelizable
- Many open questions concerning feasibility for “real world” chemistry problems
- We currently have the “integral” program, now we need to determine the wave function/density (nonlinear optimization problem)
- Our goal: QM/QM, possibly with interfacing to continuum-like models which also solve the Poisson problem at a boundary

Acknowledgement

Those who did the work...

- Luca Frediani (Dep.of Chemistry, UiT)
- Eirik Fossgaard (IRIS)
- Tor Flå (Dep.of Mathematics, UiT)

Those who paid...

- Norwegian Research Council (YFF)