

Multiresolution Quantum Chemistry in Multiwavelet Bases

Robert J. Harrison¹, George I. Fann, Takeshi Yanai, and Gregory Beylkin²

¹ Oak Ridge National Laboratory,
P.O. Box 2008 MS6367,
Oak Ridge, TN 37831-6367,
harrisonrj@ornl.gov,
<http://www.ornl.gov/~rj3>

² Department of Applied Mathematics,
University of Colorado at Boulder,
UCB 526, Boulder, CO 80309-0526,
gregory.beylkin@colorado.edu,
<http://amath.colorado.edu/faculty/beylkin/Home.html>

Abstract. Multiresolution analysis in multiwavelet bases is being investigated as an alternative computational framework for molecular electronic structure calculations. The features that make it attractive include an orthonormal basis, fast algorithms with guaranteed precision and sparse representations of many operators (e.g., Green functions). In this paper, we discuss the multiresolution formulation of quantum chemistry including application to density functional theory and developments that make practical computation in three and higher dimensions.

1 Introduction

Chemists are familiar with the range of length-scales that are present in molecular electronic wave functions. There is a sharp cusp at the nucleus; atoms have core and valence regions; molecules also possess inter-atomic bonding regions; and both atoms and molecules have a long exponential tail. By capturing the essence of molecules as being composed of perturbed atoms, the linear combination of atomic orbitals (LCAO) approximation has proven tremendously successful. Part of this success stems from effective management of the length scales present in the problem. The atomic orbital basis functions naturally incorporate both the short- and long-range behavior. However, the non-local nature of atomic wave functions, especially in high-precision calculations, gives rise to inefficiencies, non-physical density matrices, and numerical problems due to severe linear dependence. Furthermore, the atomic orbital basis sets must be carefully designed and calibrated [1]. If used out of context (e.g., in charge-fitting [2] or resolution-of-the-identity methods [3,4]) uncontrolled errors can be introduced.

Multiresolution analysis in multiwavelet bases [5,6,7,8] is an alternative approach that is potentially very attractive for computational chemistry. Our approach closely follows that described by Alpert et al. [5]. Among the notable features are

- a fully orthonormal basis with high-order convergence even in the presence of singularities,
- an algorithm for adaptive refinement,
- sparse representations of many operators including the Hamiltonian and Green functions, and
- fast algorithms with guaranteed precision for many common operations.

The high-order, orthonormal basis with adaptive refinement enables robust high-precision computation. The separation of length scales implicit in the multiresolution formulation provides sparse representation of operators which enables fast computation. Although the orthonormal multiwavelet basis may be used in standard matrix (i.e., Hamiltonian) formulations of quantum problems, it is often much more effective to reformulate the problem as an integral equation.

In the following, we use the Poisson equation to introduce the essential details of multiresolution in multiwavelet bases and how they result in fast algorithms. Subsequently, we discuss an essential component for efficient computation in many dimensions, which is accurate separated representations of integral operators. This is done in the context of an integral formulation of density functional theory (DFT). Finally, we briefly describe our prototype implementation.

2 Sparse Representations and Fast Algorithms

Much of our discussion is, for simplicity, in one dimension, but the generalization to higher dimensions is straightforward, except where otherwise noted.

We map the problem to the unit interval. This is recursively divided by two, so that on each level $n = 0, 1, 2, \dots$ there are 2^n boxes. In each box, we define basis functions that are the first k normalized Legendre polynomials inside the box and are zero outside. Smooth functions may be represented with an error $O(2^{-nk})$. By construction, the basis on a level is orthonormal and the functions associated with different boxes do not overlap each other. This last property of disjoint support is the most important advantage of multiwavelet bases. If boundaries or discontinuities are located at the nodes, high-order convergence may be maintained.

The basis at a given level of refinement is of dimension $k2^n$ and is labeled V_n^k . It satisfies

$$V_0^k \subset V_1^k \subset \dots \subset V_n^k \subset \dots \quad (1)$$

The multiwavelet subspace (W_n^k) is defined as the orthogonal complement of V_n^k in V_{n+1}^k

$$W_n^k = V_{n+1}^k - V_n^k \quad (2)$$

which then leads to an alternative decomposition of our basis

$$V_n^k = V_0^k + W_0^k + W_1^k + \dots + W_{n-1}^k. \quad (3)$$

Beylkin et al. give fast ($O(N)$) algorithms for converting between the two representations. Adaptive local refinement while maintaining a global norm-wise

error is accomplished by discarding small wavelet coefficients [5]. We presently use Alpert's definition of the multiwavelet basis [7].

Since V_n^k is the space of polynomials on level n , the first k moments vanish for functions in W_n^k (since they are by construction orthogonal to V_n^k). This property gives rise to sparse representations of smooth functions and operators. We use the Poisson equation in three dimensions

$$\nabla^2 u = -4\pi\rho \quad (4)$$

with free-space boundary conditions ($u(\infty) = u'(\infty) = 0$) as an example of how this enables us to replace iterative solution of differential equations with fast application of integral operators.

The matrix representation of the Laplacian in either the multiwavelet basis or standard discretizations is badly conditioned. Moreover, the largest eigenvalues are associated with the highest frequencies. This leads to the requirement of good preconditioners for the efficient iterative solution of differential equations. However, the Green's function for the Poisson equation is known and we may immediately write the solution as

$$u(r) = \int \frac{\rho(s)}{|r-s|} ds. \quad (5)$$

That the multiwavelet representation of this integral operator is sparse can be readily seen from the multipole expansion of $1/r$. Consider the interaction between two wavelets of order k in two boxes separated by r . Since the first k moments vanish, the interaction will decay as $r^{-(2k+1)}$. It is also necessary to consider the interaction between wavelets and polynomials, which will decay at worst as $r^{-(k+1)}$. We commonly employ wavelets of order 5–13, and, by increasing the order of the wavelets as we increase the required precision, it is never necessary to include interactions beyond the first 26 nearest neighbors of a box. This immediately results in an $O(N)$ algorithm with guaranteed precision. Also, the iterative solution of equations has been replaced with a single, fast, sparse matrix-vector product.

Some additional insight is provided by the telescoping series [5,6]

$$T_n = T_0 + (T_1 - T_0) + (T_2 - T_1) + \cdots + (T_n - T_{n-1}) \quad (6)$$

where T_n is the projection of an operator (or equivalently a multidimensional function) onto the polynomial basis at level n . This identity expresses an operator on the fine scale (level n) in terms of the operator on a coarse scale and successive differences between length scales. The representation of the operator at all levels may be dense, as is the case for $1/r$. However, if the operator is smooth at long range, then it will eventually be accurately represented at some scale, and thereafter the differences between levels will be small. With this awareness, it is often possible to compute efficiently and accurately without explicit introduction of the wavelet basis.

Alpert et al. [5] discusses a much wider range of differential and integral operators, including time dependent problems, for which efficient sparse representations may be constructed for arbitrary finite precision.

3 Integral Formulation of the Schrödinger Equation

The familiar Hamiltonian formulation of the Schrödinger equation

$$\left(-\frac{1}{2}\nabla^2 + V\right)\psi = E\psi \quad (7)$$

may be rewritten as an integral equation

$$\psi = -2(-\nabla^2 - 2E)^{-1}V\psi \quad (8)$$

where the inverse of the differential operator denotes application of the corresponding Green function. For bound states, the Green function is that of bound-state Helmholtz equation, which, in three dimensions, is given by

$$G(r, s, E) = \frac{e^{-\sqrt{-2E}|r-s|}}{4\pi|r-s|}. \quad (9)$$

This kernel is amenable to fast computation in the multiwavelet basis just as described for the Poisson kernel. Moreover, straightforward iteration of (8) will converge to the ground state, whereas an iterative, preconditioned eigensolver must be used to solve the differential form. For DFT models without Hartree-Fock exchange, the form of the integral operator includes the expected asymptotic exponential decay. To extract multiple roots from the integral equation we use deflation which is discussed elsewhere [8].

4 Practical Application of Integral Operators in Three Dimensions

Above, the application of integral operators has been informally demonstrated to be efficient, in the sense of scaling linearly with the number of boxes with significant coefficients. However, practical computation also requires a small prefactor. In three dimensions, the basis within each box on the locally finest level will be a tensor product of Legendre polynomials. The coefficients associated with the box will therefore carry three-indices. Similarly, a matrix representation of an operator will carry six indices; three for the source and three for the result. So it seems as if cost of both storing and applying the operator scales as $O(k^6)$, where k is again the order of the wavelet. This is prohibitively expensive since we commonly use wavelets of order 5–13.

Since the basis is orthonormal, the matrix representation of an integral operator with kernel (K) in the scaling function (Legendre polynomial) basis can be obtained by projection

$$r_{pp',qq',rr'}^{n,ll'} = \int dx dy dz dx' dy' dz' K(x-x', y-y', z-z') \phi_p^{nl}(x)\phi_{p'}^{nl'}(x')\phi_q^{nl}(y)\phi_{q'}^{nl'}(y')\phi_r^{nl}(z)\phi_{r'}^{nl'}(z') \quad (10)$$

where l and l' label the result and source boxes, and $\phi_i^{nl}(x)$ is the i 'th Legendre polynomial in box l on level n . If there existed a separated representation of the kernel that is accurate to our desired precision ϵ

$$K(x, y, z) = \sum_{i=1}^M \sigma_i X_i(x) Y_i(y) Z_i(z), \quad (11)$$

it is straightforward to see that each dimension may be separately applied. This reduces the cost of applying the operator to $O(Mk^4)$ and reduces the storage to $O(Mk^2)$. The computational cost may be further reduced by using singular value decomposition to exploit the low operator rank away from the singularity. The length of the expansion (M) is referred to as the *separation* rank. It is critical that this rank be as small as possible, and also that the coefficients in the expansion be small and of the same sign so as to avoid loss of precision.

Recently [9], it has been shown that many physically significant operators, including the Poisson kernel and even the entire many-electron Schrödinger Hamiltonian, have low separation rank that increases only logarithmically with both the precision and dimension. This hints at the possibility of alternative schemes for efficient computation in many dimensions. In the present context, it ensures the existence of efficient representations but it does not tell us how to construct them.

Beylkin et al. [10] have developed a powerful approach for constructing optimal separated representations of functions over finite ranges, but this is too complicated to discuss here. However, near-optimal representations may be formed from application of quadrature rules to integral representations of kernels as a sum of Gaussians

$$K(r) = \sum_{i=1}^M c_i e^{-t_i r^2}. \quad (12)$$

For instance, we have the standard identity

$$\frac{e^{-\mu r}}{r} = \frac{2}{\sqrt{\pi}} \int_0^\infty e^{-r^2 t^2 - \frac{\mu^2}{4t^2}} dt \quad (13)$$

which with the substitution $t = e^s$ can be transformed to

$$\frac{e^{-\mu r}}{r} = \frac{2}{\sqrt{\pi}} \int_{-\infty}^\infty e^{-r^2 e^{2s} - \frac{\mu^2}{4} e^{-2s} + s} ds. \quad (14)$$

The advantages of this form are that the integrand dies very rapidly at both ends of the quadrature range so the trapezoidal rule converges geometrically, and that the required resolution (bandlimit) is almost independent of r . The resulting Gaussian exponents form a geometric series ensuring a uniform error at all length scales. The form of the integrand may be analyzed to establish a suitable range and empirical tests suffice to find the required resolution for given precision. The results are

$$s_{hi} = \log(T/r_{i0}^2)/2 \quad (15)$$

$$s_{lo} = -\log(4T/\mu^2)/2 \quad (16)$$

$$1/h = .2 + 0.47 \log_{10} \epsilon \quad (17)$$

where the range of the quadrature is $[s_{lo}, s_{hi}]$, the expansion has a relative precision of ϵ over the range $[r_{lo}, 1]$, and h is the spacing of the quadrature points. The parameter T is empirically determined to be 5, 10, 14, 18, and 24, respectively for accuracies 1e-2, 1e-4, 1e-6, 1e-8, and 1e-10. However, because of the super-exponential decay, the number of quadrature points is only weakly dependent upon T , so using the optimal value of T saves at most only 1 or 2 points. The quadrature may be modified to provide compact representations of the Poisson kernel ($\mu = 0$). To illustrate the automatic procedure, fits were generated for $\exp(-30r)/r$ over $[1\text{-e}8, 1]$ for accuracies 1e-10, 1e-8, 1e-6, 1e-4 and 1e-2. The errors are plotted in figure 1.

5 Prototype Implementation

In the current implementation, multiwavelets of any order are supported. The two-scale coefficients are generated using Alpert's algorithm [7] in extended precision floating point numbers in Python (e.g., 156-bit arithmetic is used to generate the two-scale coefficients for order 10). The extended precision is necessary only to generate the coefficients which are stored for subsequent use. Standard double precision arithmetic is used for all other operations.

Our initial implementation uses Python for high-level control and C/Fortran for computationally intensive operations including matrix transformations, quadratures and the innermost loops. At the highest level, we have defined a **Function** class that includes methods for evaluation, compression, reconstruction, addition, multiplication by a function or scalar, differentiation, application of the Laplacian, and other operations. The operator overloading capabilities of Python provide great expressivity and enable very compact programs. For instance, if `psi` is an instance of the **Function** class representing an orbital, and similarly `V` represents the potential, then the following statement applies the Fock operator to the orbital¹

```
Hpsi = -0.5 * (Delsq * psi) + V*psi
```

`Delsq` is an empty class that is never instantiated. If an instance of **Function** is multiplied on the right by `Delsq`, the function's Laplacian method is invoked. Evaluation of a function at a point with the natural semantics `psi(x,y,z)` is accomplished by overloading the function call operator.

6 Conclusions

Multiresolution in multiwavelet bases is an attractive alternative framework for computation in chemistry. It provides fast algorithms with guaranteed precision.

¹ Note that matrix elements of the Laplacian, including evaluation of the kinetic energy, should employ the variational form $\langle f|\nabla^2|g \rangle = -\langle \nabla f \cdot \nabla g \rangle$.

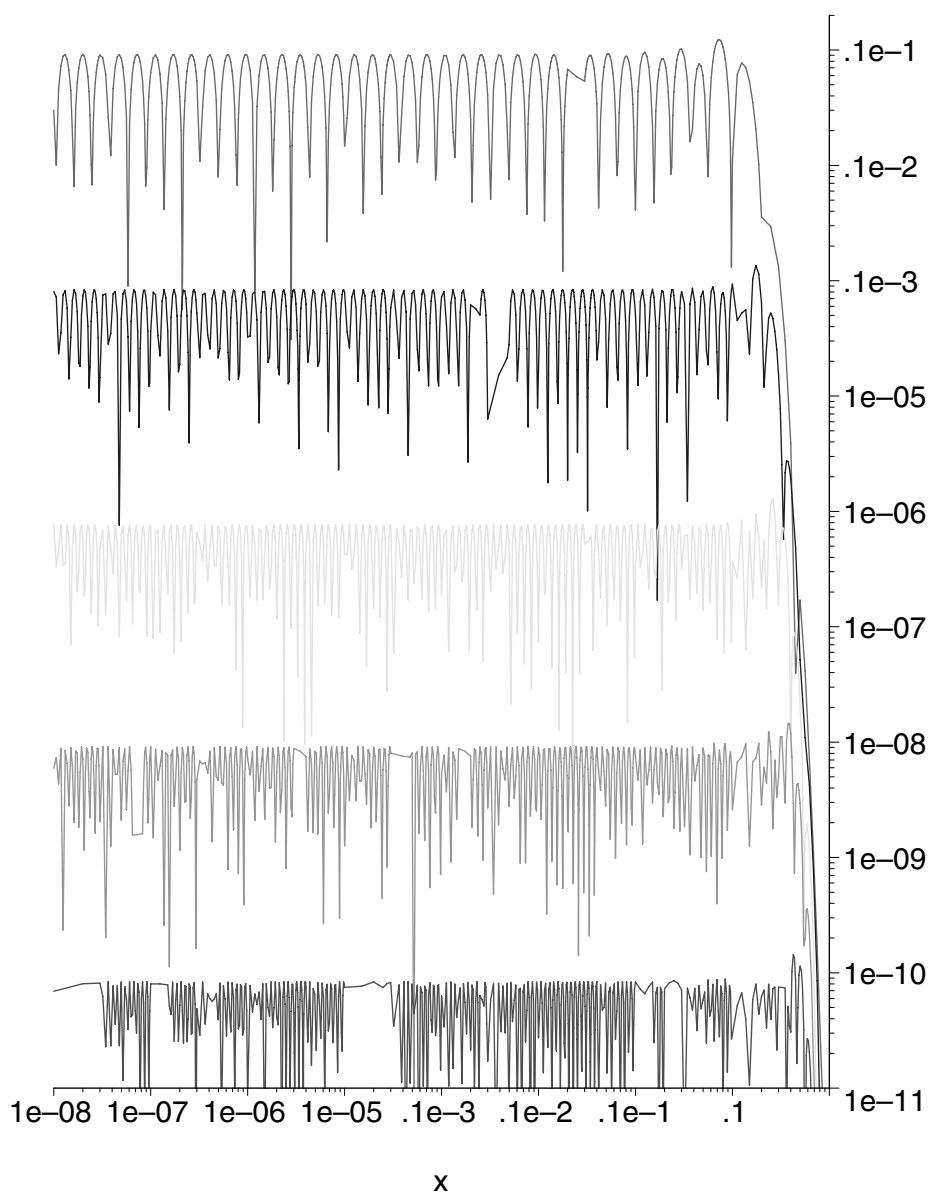


Fig. 1. Automatically generated representations of $\exp(-30r)/r$ accurate to $1e-10$, $1e-8$, $1e-6$, $1e-4$, and $1e-2$ (measured by the weighted error $r(\exp(-30r)/r - \text{fit}(r))$) for r in $[1e-8, 1]$ were formed with 92, 74, 57, 39 and 21 terms, respectively

The orthonormal basis enables all of the familiar results of quantum theory to be used without change, though the greatest benefit is obtained by reformulating the entire solution scheme to take advantage of the properties of the basis. In particular, the sparse representation of many integral operators, including many physically significant Green functions, enables new approaches. Additional techniques are necessary for scattering states for which the corresponding Green functions do not become smooth at long range. Compact separated representations are important for efficient computation in higher dimensions, and may also be of utility in conventional Gaussian-based methods. Finally, although our current implementation uses orbitals, so cannot attain full linear scaling, it is well established [11,12,13] that a density-matrix approach in a wavelet basis will achieve linear scaling while maintaining a guarantee of arbitrary, finite precision.

References

1. Dunning, Jr., T.H., Peterson, K.A., and Woon, D. E.: Gaussian Basis Sets for Use in Correlated Calculations, in Encyclopedia of Computational Chemistry, Ed., P. v. R. Schleyer, (John Wiley & Sons Ltd., 1997)
2. Dunlap, B.I., Connolly, J.W.D., and Sabin, J.R.: J. Chem. Phys. **71** (1979) 3396 and 4993
3. Vahtras, O., Almlöf, J., and Feyereisen, M. W.: Integral approximations for LCAO-SCF calculations. Chem. Phys. Lett. **213** (1993) 514; Feyereisen, M., Fitzgerald, G., and Komornicki, A.: Chem. Phys. Lett. **208** (1993) 359
4. Challacombe, M., Schwegler E., and Almlöf, J.: pp 53–107 in Review of Current Trends, ed. J. Leczszynski (World Scientific, 1996)
5. Alpert, B., Beylkin, G., Gines, D., and Vozovoi, L.: Adaptive Solution of Partial Differential Equations in Multiwavelet Bases, J. Comp. Phys. **182** (2002) 149–190
6. Beylkin, G., Coifman, R., and Rokhlin, V.: Fast wavelet transforms and numerical algorithms, I, Comm. Pure. Appl. Math. 44 **141** (1991) 183
7. Alpert, B.K.: A Class of Bases in L^2 for the Sparse Representation of Integral Operators, SIAM J. Math. Anal. **24** (1993) 246–262
8. Harrison, R.J., Fann, G.I., Yanai, T., and Beylkin, G.: Multiresolution Quantum Chemistry, in preparation, 2003.
9. Beylkin, G. and Mohlenkamp, M.J.: Numerical operator calculus in higher dimensions, Proc. Nat. Acad. Sci. **99** (2002) 10246–10251
10. Beylkin, G. and Monzon, L., work in progress.
11. Li, X.-P.; Nunes, R.W. and Vanderbilt, D.: Density-matrix electronic-structure method with linear system-size scaling, Phy. Rev. B **47** (1993) 10891
12. Beylkin, G., Coult, N. and Mohlenkamp, M.J.: Fast Spectral Projection Algorithms for Density-Matrix Computations, J. Comp.Phys. **152** (1999) 32–54
13. Goedecker, S.: Linear Scaling Electronic Structure Methods, Rev. Mod. Phys. **71** (1999) 1085