Algorithms for Numerical Analysis in High Dimensions.
Part II: Operators in high dimensions

Gregory Beylkin
with
M. Mohlenkamp (Ohio U.) , L. Monzón, F. Pérez, V. Cheruvu (CU), R. Cramer (now at Raytheon),
and R. Harrison, G. Fann, T. Yanai, & Z. Gan (ORNL)

IPAM
October 19, 2004

Plan of the talk

Separated representations for:

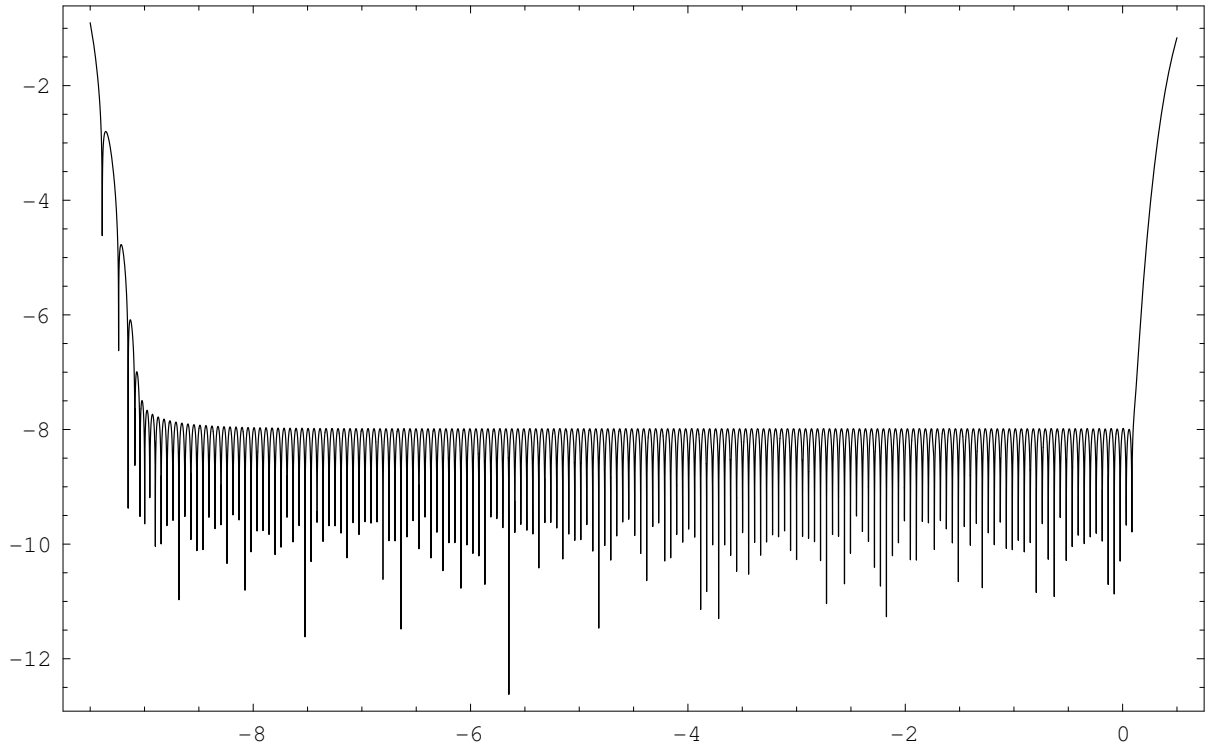
1. Green's functions
2. Lattice sums
3. Green's function for the confining harmonic potential
4. 3D applications: electron structure computations in quantum chemistry
5. Fast Poisson solver
6. Schrödinger operator,
7. Antisymmetric wave function

An example of separated representation

We have

$$\left| \frac{1}{||r||} - \sum_{m=1}^{M} w_m e^{-p_m ||r||^2} \right| \leq \frac{\epsilon}{||r||},$$

for $\delta \leq ||r|| \leq 1$, where $p_m, w_m > 0$ and $M = \mathcal{O}(-\log \delta)$.



Error ($\log_{10}$) of approximating the Poisson kernel for $10^{-9} \leq ||r|| \leq 1$, $M = 89$.

Separated representations via quadratures

We use the trapezoidal rule to discretize the integral

$$r^{-\alpha} = \frac{2}{\Gamma(\alpha/2)} \int_{-\infty}^{\infty} e^{-r^2 e^{2s} + \alpha s} ds$$

to obtain the initial approximation.

Theorem 1. *For any $\alpha > 0$, $0 < \delta \leq 1$, and $0 < \epsilon \leq \min\left\{\frac{1}{2}, \frac{8}{\alpha}\right\}$, there exist positive numbers p_m and w_m such that*

$$\left| r^{-\alpha} - \sum_{m=1}^{M} w_m e^{-p_m r^2} \right| \leq r^{-\alpha}\epsilon, \ \textit{for all } \delta \leq r \leq 1$$

with

$$M = \log \epsilon^{-1}[c_0 + c_1 \log \epsilon^{-1} + c_2 \log \delta^{-1}],$$

where c_k are constants that only depend on α. For fixed power α and accuracy ϵ, we have $M = \mathcal{O}(\log \delta^{-1})$.

Nearly optimal quadratures

The number of nodes obtained via Theorem 1 is of correct order, but not close to optimum (mostly in the range of small exponents, away from the singularity).

We optimize the quadrature by solving the following approximation problem: given

$$f(x) = \sum_{m=1}^{M_0} b_m \mathrm{e}^{-\tau_m x},$$

and $\epsilon > 0$, find a function (of the same form),

$$g(x) = \sum_{m=1}^{M} w_m \mathrm{e}^{-t_m x},$$

with $M < M_0$ and such that

$$|f(x) - g(x)| \leq \epsilon, \text{ for } x \in [0, 1].$$

A connection to optimal rational approximations

- Our method (developed with Lucas Monzón) uses finite Hankel matrices.
- The nearly optimal nodes are found as zeros of certain functions constructed from the singular vectors of these matrices.
- If approximation were on the interval $[0, \infty]$ then, by using the Laplace transform, the problem reduces to that for rational functions.
- Finding best L^∞ approximation for rational functions is a part of Arov, Adamjan, Krein theory (circa 1970).

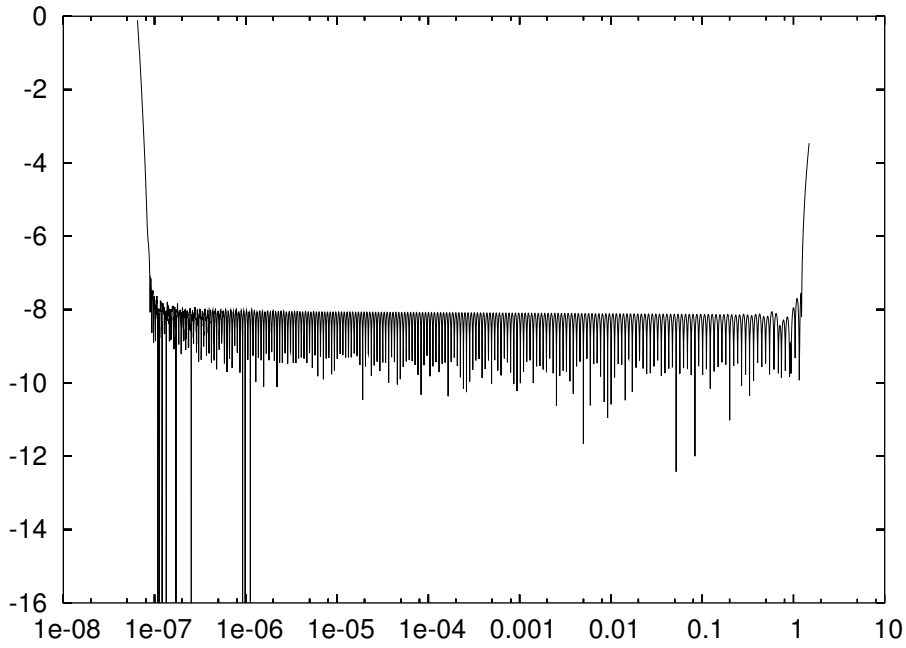
Not enough time to go into details

The projector on the divergence-free functions

For the projector on the divergence-free functions, we have

$$\left| \frac{1}{||r||^3} - \sum_{m=1}^{M} w_m e^{-p_m ||r||^2} \right| \leq \frac{\epsilon}{||r||^2},$$

for $\delta \leq ||r|| \leq 1$, where $p_m, w_m > 0$ and $M = \mathcal{O}(-\log \delta)$.



Error of the approximation with 110 terms over the domain $10^{-7} \leq ||r|| \leq 1$.

How do we use it?

We need a few additional constructions ...

Multiresolution Analysis

Chain of subspaces:

$$\ldots \subset V_0 \subset V_1 \subset V_2 \subset \ldots$$

$\bigcap_j V_j = \{0\}$ and $\overline{\bigcup_j V_j} = L^2(R^d)$.

Examples: piecewise-constant functions, Daubechies' scaling functions, polynomials up to degree $\mathfrak{m} - 1$ on a collection of intervals.

Detail spaces W_j : $V_{j+1} = V_j \oplus W_j$

Orthonormal bases in V_j and W_j are defined by the scaling function ϕ and wavelet ψ.

Examples: Haar basis, Daubechies' wavelets, multiwavelets.

Projectors: P_j onto V_j, Q_j onto W_j.

Non-standard representation of operators

Consider

$T : L_2 \to L_2$

Define

$T_j = P_j T P_j, \quad A_j = Q_j T Q_j, \quad B_j = Q_j T P_j \quad C_j = P_j T Q_j,$ for $j \in Z.$

We have the telescopic expansion,

$$T_n - T_0 = \sum_{j=1}^{j=n} (P_j T P_j - P_{j-1} T P_{j-1}) = \sum_{j=1}^{j=n} (T_j - T_{j-1}).$$

Choice of bases

We use multiwavelet bases

On each scale the scaling functions are orthogonal polynomials of degree up to $\mathfrak{m} - 1$ on subintervals.

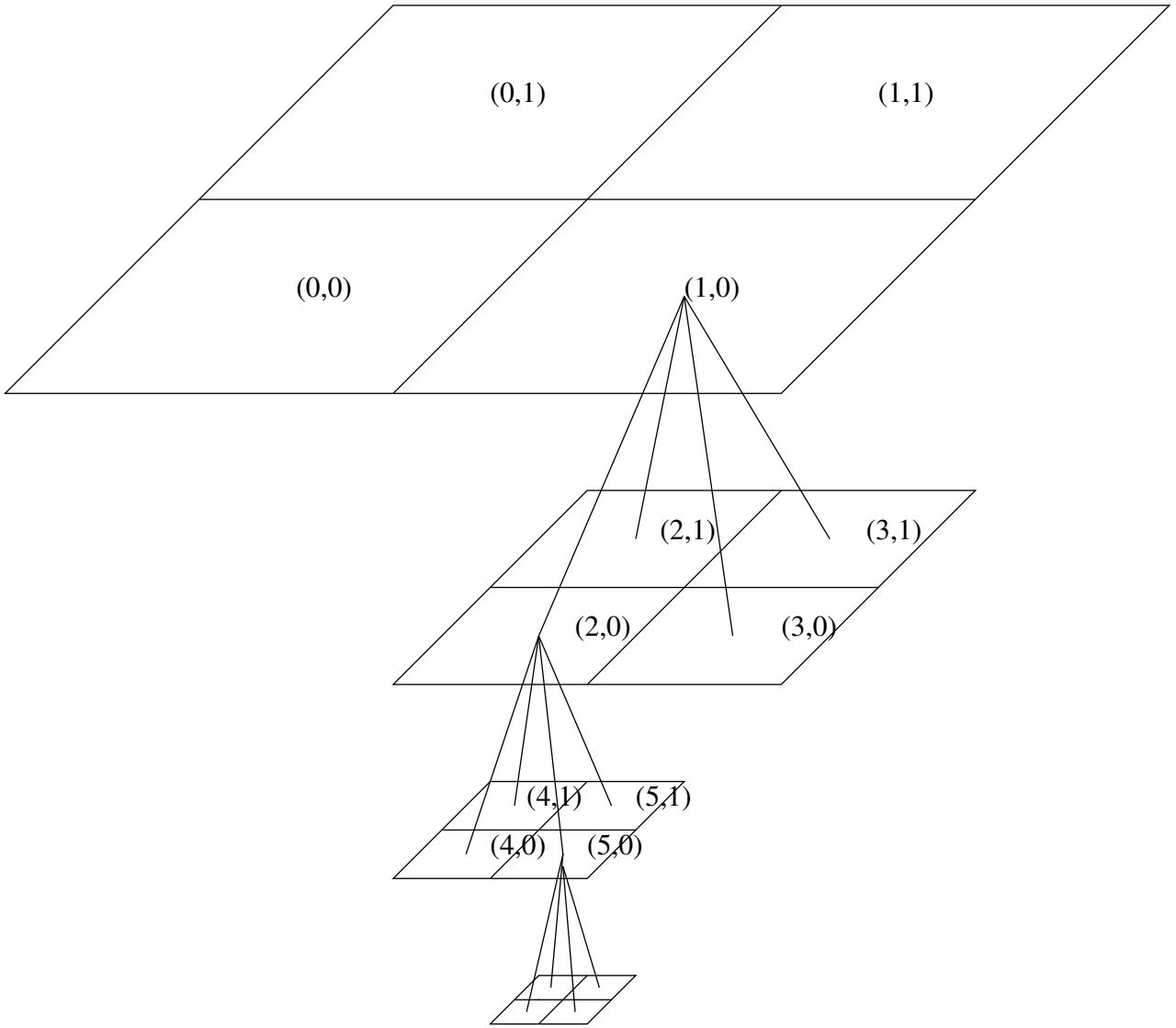
Choices:

1. The Legendre polynomials

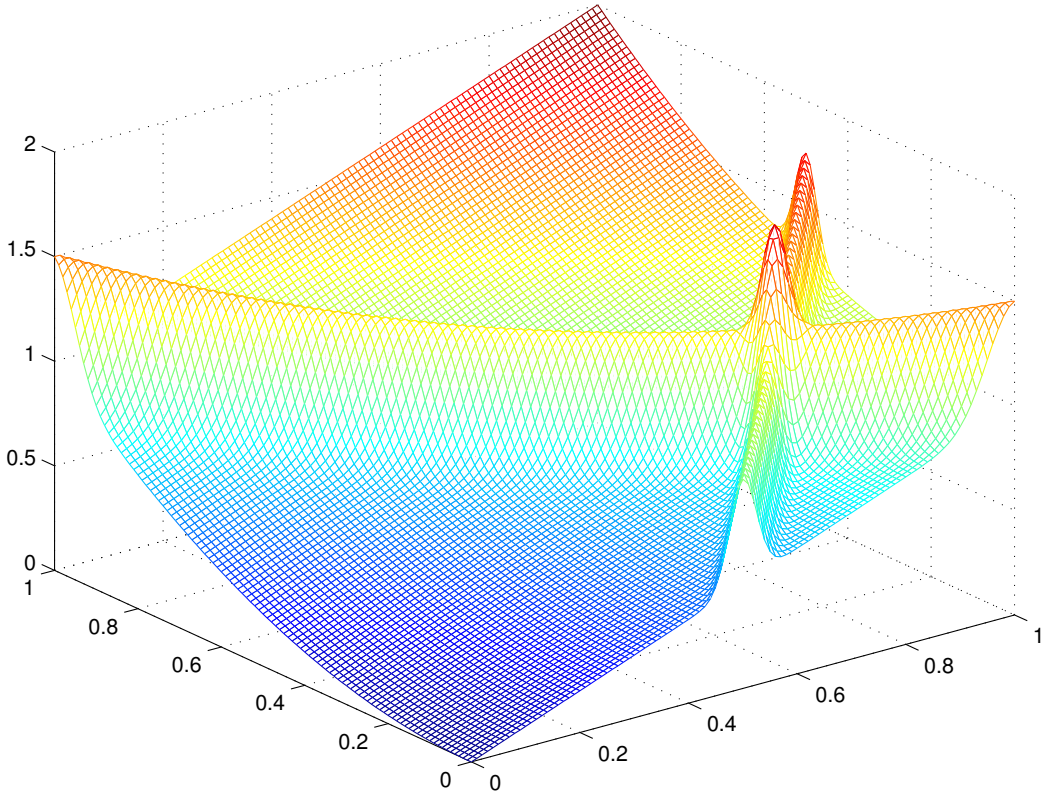
2. The Lagrange interpolating polynomials with the Legendre nodes

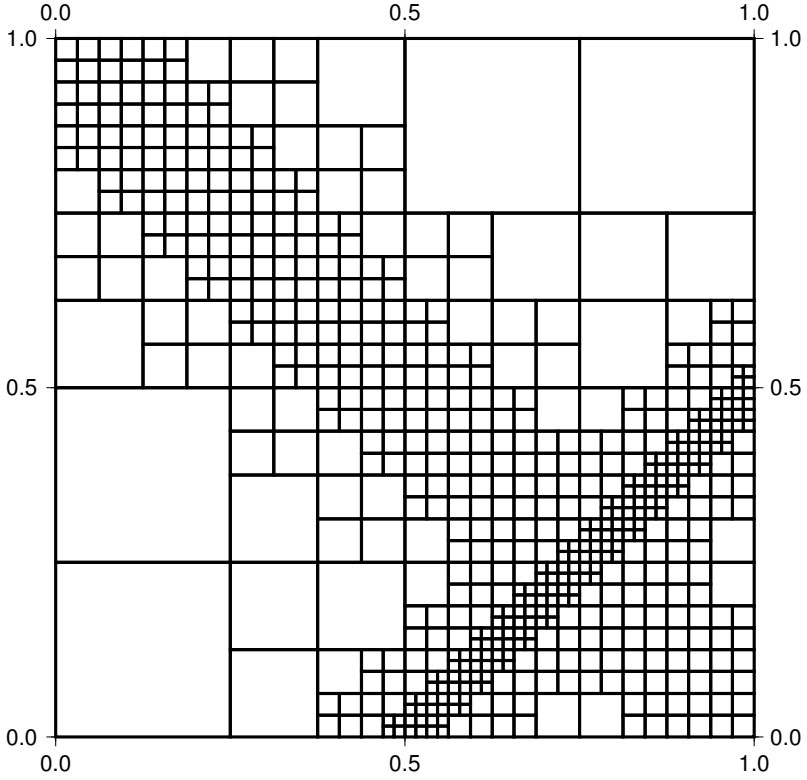
Many useful properties

Adaptive subdivision



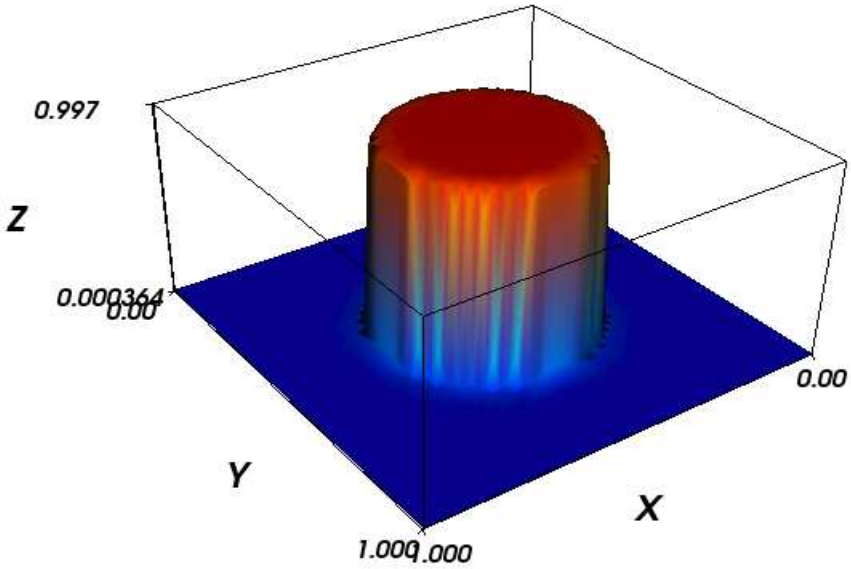
Adaptive representation of a function



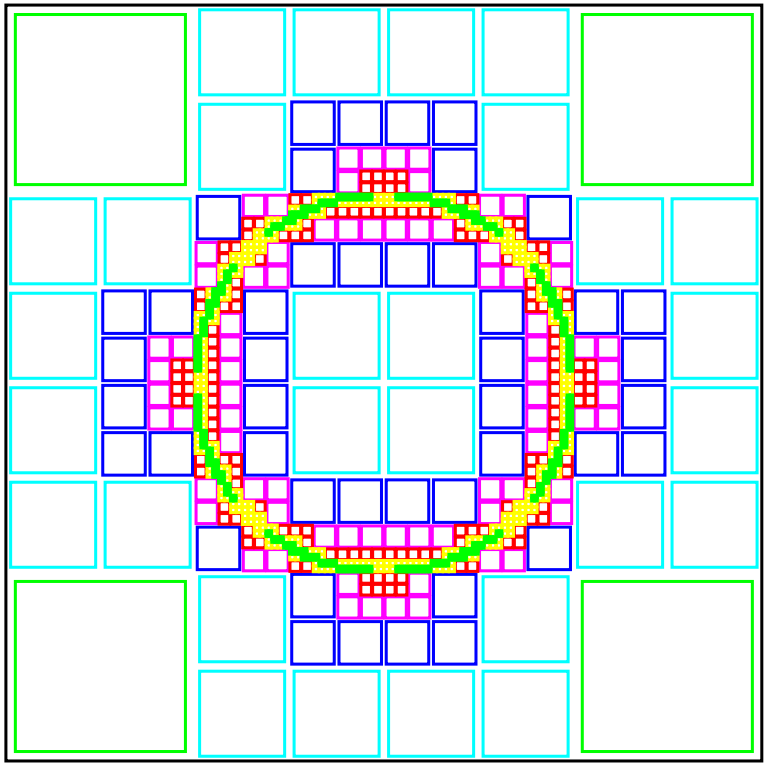


Another example

Consider the characteristic function of a disk



$N_{nod} = 8, \epsilon = 1.0e-02, N_{blocks} = 1276$



The cross-correlation functions of scaling functions

For convolution operators we only need the cross-correlation functions of the scaling functions, namely,

$$\Phi_{ii'}(x) = \begin{cases} \Phi^{+}_{ii'}(x), & 0 \le x \le 1, \\ \Phi^{-}_{ii'}(x), & -1 \le x < 0, \\ 0, & 1 < |x|, \end{cases}$$

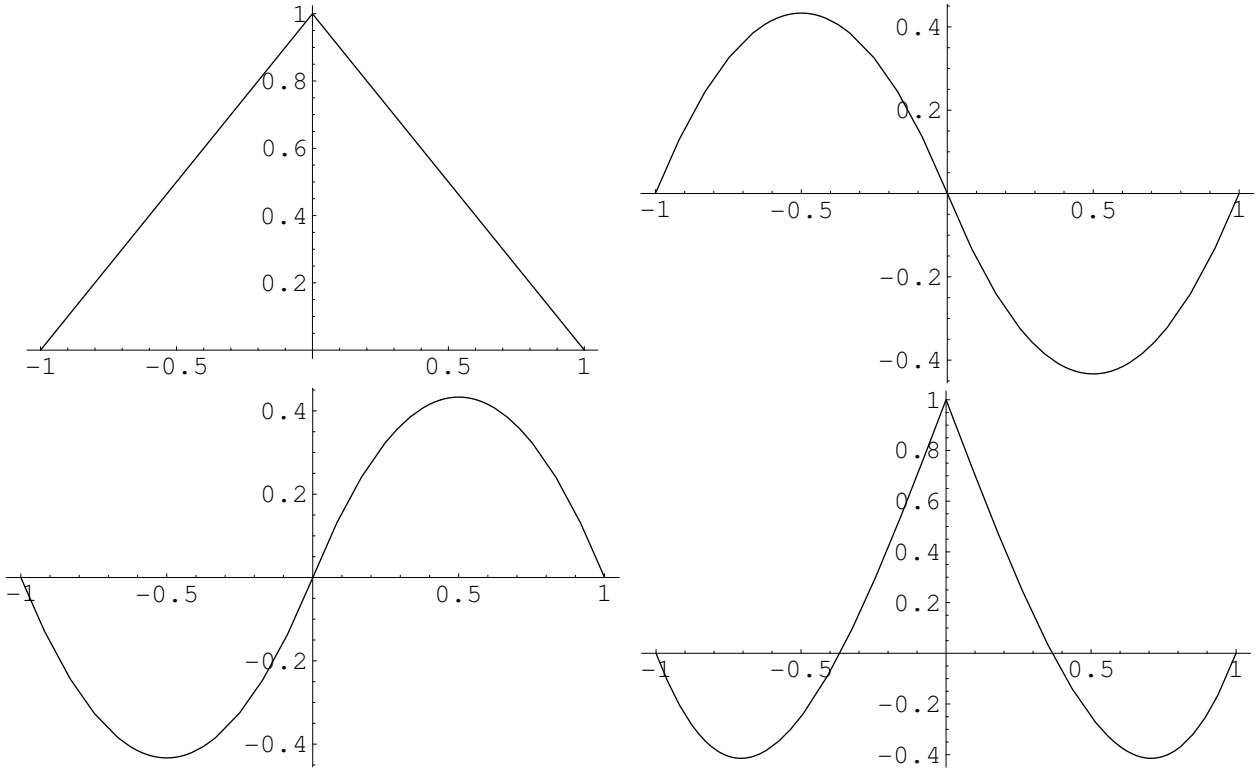
where $i, i' = 0, \ldots, \mathfrak{m}-1$, $\mathfrak{m}$ is the order of the basis, and

$$\Phi^{+}_{ii'}(x) = \int_0^{1-x} \phi_i(x+y)\,\phi_{i'}(y)dy\,, \quad \Phi^{-}_{ii'}(x) = \int_{-x}^{0} \phi_i(x+y)\,\phi_{i'}(y)dy\,.$$

The scaling functions ϕ_i are the normalized Legendre polynomials on the interval $[0,1]$,

$$\phi_i(x) = \begin{cases} \sqrt{2i+1}P_i(2x-1), & x \in [0,1] \\ 0, & x \notin [0,1] \end{cases},$$

where P_i are the Legendre polynomials on $[-1,1]$. This implies that the functions $\Phi_{ii'}$ are piecewise polynomials of degree $i+i'+1$ with the support in $[-1,1]$.



The first four cross-correlation functions $\Phi_{00}, \Phi_{01}, \Phi_{10}$ and Φ_{11}.

The Poisson kernel

Due to the homogenuity of the Poisson kernel, we have

$$t^{n;\,\mathbf{l}}_{ii',jj',kk'} = 2^{-2n}\, t^{\mathbf{l}}_{ii',jj',kk'} \ ,$$

where

$$t^{\mathbf{l}}_{ii',jj',kk'} = t^{l_1,l_2,l_3}_{ii',jj',kk'} = \frac{1}{4\pi} \int_{-1}^{1}\int_{-1}^{1}\int_{-1}^{1} \frac{1}{||\mathbf{x}+\mathbf{l}||}\, \Phi_{ii'}(x_1)\, \Phi_{jj'}(x_2)\, \Phi_{kk'}(x_3)\, d\mathbf{x},$$

and

$$\Phi_{ii'}(x) \ = \ \int_0^1 \phi_i(x+y)\, \phi_{i'}(y) dy \, , \quad i,i' = 0,\ldots,k-1 \, ,$$

are the cross-correlation functions of the scaling functions of the multiwavelet basis.

Separated representation of the Poisson kernel

Theorem: For any $\epsilon > 0$ the coefficients $t^{\mathbf{l}}_{ii',jj',kk'}$ have an approximation with a low separation rank,

$$r^{\mathbf{l}}_{ii',jj',kk'} = \sum_{m=1}^{M} \frac{w_m}{b} F^{m,l_1}_{ii'} F^{m,l_2}_{jj'} F^{m,l_3}_{kk'} ,$$

such that

$$\begin{aligned} |t^{\mathbf{l}}_{ii',jj',kk'} - r^{\mathbf{l}}_{ii',jj',kk'}| &\leq \frac{2\epsilon}{\pi} && \max_i |l_i| \geq 2 \\ |t^{\mathbf{l}}_{ii',jj',kk'} - r^{\mathbf{l}}_{ii',jj',kk'}| &\leq C\delta^2 + \frac{2\epsilon}{\pi} && \max_i |l_i| \leq 1 \end{aligned}$$

$$F^{m,l}_{ii'} = \int_{-1}^{1} e^{-p_m/b^2(x+l)^2} \Phi_{ii'}(x)\, dx ,$$

$b = \sqrt{3} + ||\mathbf{l}||$, and δ, $M = O(-\log \delta) + O(-\log \epsilon)$, p_m, w_m, $m = 1, \ldots, M$ are from the separated representation of the kernel.

Other important examples

1. The projector on the divergence-free functions (similar to the Poisson kernel)

2. The Helmholtz kernel, $e^{iKr}/r \approx \sum_m w_m e^{\tau_m r^2}$ for $K \sim 100$ (some other time)

3. Lattice Sums

4. The Green's function for (N-particle) confining harmonic potential

$$\mathcal{H} = \sum_{j=1}^{N} (-\Delta_j + ||\mathbf{x}_j||^2)$$

$$\mathcal{H}^{-1}(\mathbf{x}, \mathbf{y}) = \sum_{m=1}^{M} \hat{w}_m \prod_{j=1}^{N} e^{-\tau_m ||\mathbf{x}_j||^2} e^{-\sigma_m ||\mathbf{x}_j - \mathbf{y}_j||^2} e^{-\tau_m ||\mathbf{y}_j||^2}$$

Separated representations for computing lattice sums

- Lattice sums and periodic Green's functions appear in many applications
- Examples: lattice sums for gratings, elastodynamics, optics, chemistry ...
- Many publications, mostly using the spherical harmonics, summation formulas ... McPhedran, Movchan, ...
- Variants of Ewald summation remain methods of choice
- Huang's integral representations of harmonic lattice sums

Lattice sums for Green's Functions

Formally, the periodic Green's function for the Poisson equation is given by

$$G(\mathbf{x}) = \sum_{\mathbf{n}\in\mathbb{Z}^3} \frac{1}{||\mathbf{x}+\mathbf{n}||} = \frac{1}{||\mathbf{x}||} + \sum_{\mathbf{n}\in\mathbb{Z}^3}{}' \frac{1}{||\mathbf{x}+\mathbf{n}||},$$

for $\mathbf{x} \in B$, where $B = [-1/2, 1/2]^3$ and $\mathbf{n} = (n_1, n_2, n_3)$.

The Green's function with the zero b.c. is given by

$$G_0(\mathbf{x}) = \sum_{\mathbf{n}\in\mathbb{Z}^3} \frac{(-1)^{n_1+n_2+n_3}}{||\mathbf{x}+\mathbf{n}||} = \frac{1}{||\mathbf{x}||} + \sum_{\mathbf{n}\in\mathbb{Z}^3}{}' \frac{(-1)^{n_1+n_2+n_3}}{||\mathbf{x}+\mathbf{n}||}.$$

Approximation of lattice sums

Rescaling, we obtain for $1/2 \leq ||\mathbf{x}|| \leq 1/(2\delta)$

$$\Big| \frac{1}{||\mathbf{x}||} - \sum_{m=1}^{M} (2\delta w_m) e^{-4p_m \delta^2 ||\mathbf{x}||^2} \Big| \leq \frac{\epsilon}{||\mathbf{x}||}.$$

We set $\rho_m = 2\delta w_m$ and $t_m = 4p_m\delta^2$.

Assuming that we assign the lattice sum Σ' a finite value, we have for $\mathbf{x} \in B$,

$$| \sum_{\mathbf{n}\in\mathbb{Z}^3}{}' \frac{1}{||\mathbf{x}+\mathbf{n}||} - \sum_{m=1}^{M} \rho_m \sum_{\mathbf{n}\in\mathbb{Z}^3}{}' e^{-t_m||\mathbf{x}+\mathbf{n}||^2} || \leq \epsilon \sum_{\mathbf{n}\in\mathbb{Z}^3}{}' \frac{1}{||\mathbf{x}+\mathbf{n}||}.$$

Approximation of the Green's function

We obtain an approximation to G as

$$\frac{1}{||\mathbf{x}||} - \sum_{m=1}^{M} \rho_m e^{-t_m ||\mathbf{x}||^2} + G^{per}(\mathbf{x}),$$

where G^{per} is the periodic component,

$$G^{per}(\mathbf{x}) = \sum_{m=1}^{M} \rho_m \sum_{\mathbf{n} \in \mathbb{Z}^3} e^{-t_m ||\mathbf{x}+\mathbf{n}||^2}.$$

The combination

$$\frac{1}{||\mathbf{x}||} - \sum_{m=1}^{M} \rho_m e^{-t_m ||\mathbf{x}||^2}$$

is less than ϵ outside the ball of radius $1/2$.

Smooth periodic component

For the periodic function G^{per}, we have

$$G^{per}(\mathbf{x}) = \sum_{\mathbf{p}\in\mathbb{Z}^3} \hat{g}_{\mathbf{p}} e^{2\pi i \mathbf{x}\cdot\mathbf{p}},$$

where

$$\hat{g}_{\mathbf{p}} = \pi^{3/2} \sum_{m=1}^{M} \frac{\rho_m}{t_m^{3/2}} e^{-\pi^2\mathbf{p}^2/t_m}, \qquad \mathbf{p} \neq \mathbf{0}$$

and $\hat{g}_{\mathbf{0}} = 0$.

A Madelung-Like sum

As an example, compute

$$S = \sum_{\mathbf{n}\in\mathbb{Z}^3}{}' \frac{(-1)^{n_1+n_2+n_3}}{||\mathbf{n}||},$$

which according to D. Borwein, J.M. Borwein, C. Pinner, in "Convergence of Madelung-Like sums", Trans. Amer. Math. Soc., v. 350, 8, 1998,

$$S = -1.74756459...$$

We have

$$S = \pi^{3/2} \sum_{m=1}^{M} \rho_m [\sum_{n\in\mathbb{Z}} \frac{1}{\sqrt{t_m}} e^{-(n+1/2)^2\pi^2/t_m}]^3 - \sum_{m=1}^{M} \rho_m = -1.7475645946...$$

A multiresolution approach

1. Given the free-space operator, construct its telescopic series

$$T = T_0 + (T_1 - T_0) + (T_2 - T_1) + \dots$$

2. Compute lattice sums on each scale separately; they are unconditionally convergent except that for T_0.

3. The projection T_0 of the Poisson kernel is given by the integrals

$$t^{\mathbf{n}}_{ii',jj',kk'} = \int_B \frac{1}{||\mathbf{x}+\mathbf{n}||}\,\Phi_{ii'}(x_1)\,\Phi_{jj'}(x_2)\,\Phi_{kk'}(x_3)\,d\mathbf{x}\,,$$

where $\mathbf{n} \in \mathbb{Z}^3$, $B = [-1/2, 1/2]^3$, $\mathbf{x} = (x_1, x_2, x_3)$, $i, i', j, j', k, k' = 0, \dots, \mathfrak{m}-1$ and $\mathfrak{m}$ is the order of the basis.

Lattice sum for T_0

We need to define lattice sums of *integrals* with cross-correlation functions

$$t^{per}_{ii',jj',kk'} = \sum_{\mathbf{n}\in\mathbb{Z}^3} t^{\mathbf{n}}_{ii',jj',kk'} = \sum_{\mathbf{n}\in\mathbb{Z}^3} \int_B \frac{1}{||\mathbf{x}+\mathbf{n}||}\,\Phi_{ii'}(x_1)\,\Phi_{jj'}(x_2)\,\Phi_{kk'}(x_3)\,d\mathbf{x}\,.$$

Let us consider $t^{\mathbf{n}}_{ii',jj',kk'}$ as a function of $\mathbf{n}\in\mathbb{Z}^3$. Using

$$\frac{1}{||\mathbf{x}+\mathbf{n}||} = \frac{1}{||\mathbf{n}||} - \frac{\mathbf{x}\cdot\mathbf{n}}{||\mathbf{n}||^3} - \frac{1}{2}\frac{||\mathbf{x}||^2}{||\mathbf{n}||^3} + \frac{3}{2}\frac{(\mathbf{x}\cdot\mathbf{n})^2}{||\mathbf{n}||^5} + \mathcal{O}(\frac{1}{||\mathbf{n}||^4}),$$

we observe that the coefficient $t^{per}_{00,00,00}$ cannot be given any meaning.

But if we apply the resulting operator to periodic functions with zero mean, the coefficient $t^{per}_{00,00,00}$ is not needed and we simply set $t^{per}_{00,00,00} = 0$.

Convergence due to vanishing moments

For $i + i' \geq 1$ all functions $\Phi_{ii'}$ have vanishing moments, namely,

$$\int_{-1}^{1} \Phi_{ii'}(x)\, x^m\, dx = 0, \quad 0 \leq m \leq i + i' - 1.$$

Integrals involving functions $\Phi_{ii'}$ with indices $i + i' \geq 3$ yield a rapid decay of the coefficients,

$$t^{\mathbf{n}}_{ii',jj',kk'} = \mathcal{O}(\frac{1}{||\mathbf{n}||^4}),$$

and the absolute convergence of the corresponding lattice sums.

Conditional convergence

For functions $\Phi_{ii'}$ with indices $1 \leq i + i' \leq 2$, we have (for example)

$$t_{01,00,00}(\mathbf{n}) = \frac{n_1}{||\mathbf{n}||^3} \left(\int_{-1}^{1} x \, \Phi_{01}(x) dx \right) \left(\int_{-1}^{1} \Phi_{00}(x) dx \right)^2 + \mathcal{O}(\frac{1}{||\mathbf{n}||^4}),$$

or

$$t_{11,00,00}(\mathbf{n}) = \frac{2n_1^2 - n_2^2 - n_3^2}{2||\mathbf{n}||^5} \left(\int_{-1}^{1} x^2 \, \Phi_{11}(x) dx \right) \left(\int_{-1}^{1} \Phi_{00}(x) dx \right)^2 + \mathcal{O}(\frac{1}{||\mathbf{n}||^4}).$$

Summing over appropriately chosen domains of indices, we have conditional convergence.

The Green's Function for N-particle confining harmonic potential

First we compute the kernel of the operator $e^{-t\mathcal{H}_1}$, where

$$\mathcal{H}_1 = -\frac{d^2}{dx^2} + x^2$$

is the Hamiltonian for the confining harmonic potential in 1D.

$\mathcal{H}_1$ has discrete spectrum $\lambda_n = 2n + 1$, $n = 0, 1, \ldots$, and its eigenfunctions are well-known so that

$$K_{\mathcal{H}_1}(x, y) = e^{-(x^2+y^2)/2} \sum_{n=0}^{\infty} \frac{\lambda_n}{\sqrt{\pi} 2^n n!} H_n(x) H_n(y),$$

where H_n are the Hermite polynomials.

The kernel of $e^{-t\mathcal{H}_1}$

For the kernel of $e^{-t\mathcal{H}_1}$, we have

$$K_{e^{-t\mathcal{H}_1}}(x,y) = e^{-(x^2+y^2)/2} \sum_{n=0}^{\infty} \frac{e^{-t\lambda_n}}{\sqrt{\pi} 2^n n!} H_n(x) H_n(y),$$

and (with a little bit of work)

$$K_{e^{-t\mathcal{H}_1}}(x,y) = \frac{1}{\sqrt{2\pi \sinh(2t)}} e^{-\frac{(x-y)^2}{2\sinh(2t)}} e^{-\tanh(t)\,(x^2+y^2)/2}.$$

Spectral approximation of $\mathcal{H}_1^{-1}$.

We approximate the kernel of the Green's function $\mathcal{G} = \mathcal{H}_1^{-1}$,

$$G(x,y) = e^{-(x^2+y^2)/2} \sum_{n=0}^{\infty} \frac{1}{\lambda_n \sqrt{\pi} 2^n n!} H_n(x) H_n(y),$$

with

$$G_{\bar{n}}(x,y) = e^{-(x^2+y^2)/2} \sum_{n=0}^{\bar{n}} \frac{1}{\lambda_n \sqrt{\pi} 2^n n!} H_n(x) H_n(y),$$

so that $||\mathcal{G} - \mathcal{G}_{\bar{n}}||_2 \leq 1/\lambda_{\bar{n}}$.

Separated approximation of $\mathcal{H}_1^{-1}$.

We have $\mathcal{H}_1^{-1} = \int_0^\infty e^{-t\mathcal{H}_1}\, dt$. Since $1/\lambda_n = \int_0^\infty e^{-t\lambda_n} dt$, we approximate for $n \leq \bar{n}$

$$\Big| \int_0^\infty e^{-t\lambda_n} dt - \sum_{m=0}^{M} w_m e^{-t_m \lambda_n} \Big| \leq \epsilon,$$

where $M = \mathcal{O}(\log \bar{n})$.

Choosing $\bar{n}$ so that $1/\lambda_{\bar{n}} \leq \epsilon$, (thus, $M = \mathcal{O}(\log \epsilon^{-1})$), we arrive at the representation

$$\mathcal{H}^{-1}(\mathbf{x}, \mathbf{y}) = \sum_{m=1}^{M} \hat{w}_m \prod_{j=1}^{N} e^{-\tau_m ||\mathbf{x}_j||^2} e^{-\sigma_m ||\mathbf{x}_j - \mathbf{y}_j||^2} e^{-\tau_m ||\mathbf{y}_j||^2}$$

which is valid with accuracy ϵ in operator norm.

Importance of this example

Conjecture:

The Green's functions of other confining potentials have representations of this form, where the exponents and coefficients are determined numerically.

- The linear problem (where the exponents are given) is likely to be badly conditioned
- The non-linear problem appears to be well conditioned
- Many interesting harmonic analysis problems

Quantum Chemistry

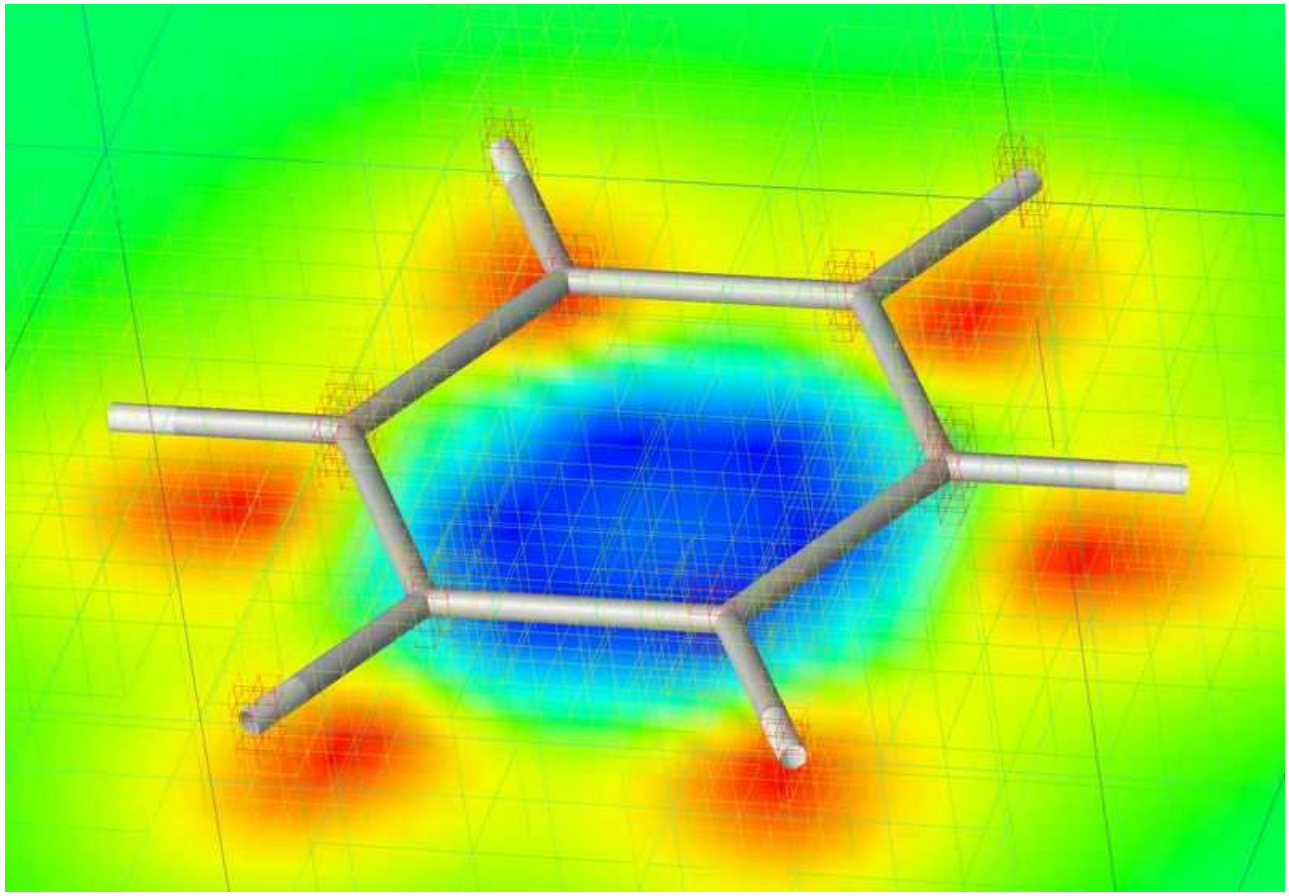
(R. Harrison, G. Fann, T. Yanai and Z. Gan (ORNL))

- Complete elimination of the basis error
- Implementation for one-electron models (HF, DFT)
- Most accurate computations up to now (within these models)
- Correct scaling of cost with system size
- Much smaller computer code than "Gaussians" (<— R. Harrison)

(R. Cramer, V. Cheruvu and F. Pérez)

- Adaptive PDE solvers
- Operator calculus in 3D

Examples: elements, small molecules...



Adaptive subdivision of space for the benzene molecule C_6H_6
(from R. Harrison, G. Fann and G. Beylkin)

Timing results for 3D adaptive Poisson solver

We compute

$$\phi(\mathbf{r}) = -\frac{1}{4\pi}\int_{\mathbb{R}^3}\frac{\rho(\mathbf{r}')}{||\mathbf{r}-\mathbf{r}'||}d\mathbf{r}$$

with ρ which gives a sum of Gaussian solutions:

$$\rho(\mathbf{r}) = -\sum_{i=1}^{3}(4\alpha^2\left|\mathbf{r}-\mathbf{r}_i\right|^2 - 2n\alpha)e^{-\alpha|\mathbf{r}-\mathbf{r}_i|^2}.$$

Timings with:

$\alpha = 300,$
$r_1 = (0.5, 0.5, 0.5),$
$r_2 = (0.6, 0.6, 0.5),$
$r_3 = (0.35, 0.6, 0.5).$

Timing results for 3D adaptive Poisson solver

Platform: Pentium 3-1.1 GHz with 512 MB of RAM

	$k = 8$	
ϵ	t_{apply} (s)	MFLOPS
10^{-3}	34	219
10^{-4}	57	227

	$k = 12$	
ϵ	t_{apply} (s)	MFLOPS
10^{-3}	18	364
10^{-4}	33	385
10^{-5}	50	390

Comments

- MADNESS (R. Harrison et. al, ORNL) is faster (but we are very close right now). Both codes are based on the same mathematics but differ algorithmically (our code is adaptive). We shoud reach and, hopefully, exceed the speed of MADNESS code.

- Already competitive with multigrid codes, especially for small ϵ, where multigrid slows down dramatically.

- Need to test against FMM-adaptive results in 2d (Greengard & Etheridge 2001).

- Python overhead is measured at a small percentage of the total time.

The Separated Representation

The standard separation of variables: $f(x_1, x_2, \ldots, x_d) = \phi_1(x_1) \cdot \phi_2(x_2) \cdot \ldots \cdot \phi_d(x_d)$

Definition: For a given ϵ, we represent a matrix $\mathbb{A} = A(j_1, j_1'; j_2, j_2'; \ldots; j_d, j_d')$ in dimension d as

$$\sum_{l=1}^{r} s_l \, A_1^l(j_1, j_1') \, A_2^l(j_2, j_2') \, \cdots A_d^l(j_d, j_d'),$$

where s_l is a scalar, $s_1 \geq \cdots \geq s_r > 0$, and $\mathbb{A}_i^l$ are matrices with entries $A_i^l(j_i, j_i')$ and norm one. We require the error to be less than ϵ:

$$||\mathbb{A} - \sum_{l=1}^{r} s_l \, \mathbb{A}_1^l \otimes \mathbb{A}_2^l \otimes \cdots \otimes \mathbb{A}_d^l|| \leq \epsilon.$$

We call the scalars s_l separation values and the rank r the separation rank.

The smallest r that yields such a representation for a given ϵ is the optimal separation rank.

Condition number of separated representation

Definition:

We call the ratio

$$\kappa = \frac{\sum_{l=1}^{r} s_l}{||\mathbb{A}||},$$

the condition number of separated representation.

It is a natural definition since $||s_l \, \mathbb{A}_1^l \otimes \mathbb{A}_2^l \otimes \cdots \otimes \mathbb{A}_d^l|| = s_l$.

We need to maintain $\kappa\mu||\mathbb{A}|| \leq \epsilon$, where μ is the machine roundoff.

Similarities and differences with SVD

If $d = 2$ then the separated representation can be obtained via SVD.

(Actually, we use a much simpler algorithm since we do not insist on orthogonality between vectors in a given direction).

If $d \geq 3$ then the analogy with SVD breaks down: by changing ϵ we change all terms in the representation rather than add/delete terms

Many attempts to treat separated representation as a generalization of SVD but the construction depends on ϵ !

Multiparticle Schrödinger operator

The Hamiltonian for the multiparticle Schrödinger operator is the sum of three terms

$$\mathcal{H} = -\sum_{i=1}^{N} \Delta_i - \sum_{i=1}^{N} V_i + \sum_{i=1}^{N-1} \sum_{m=i+1}^{N} W_{im},$$

where the 3D Laplacian corresponding to electron i is defined as $\frac{\partial^2}{\partial x_i^2} + \frac{\partial^2}{\partial y_i^2} + \frac{\partial^2}{\partial z_i^2}$, the nuclear potential V_i is operator of multiplication by $1/r_i$ and the electron-electron potential W_{im} is multiplication by $1/|r_i - r_m|$.

$\mathcal{T} = -\Delta_1 \otimes \mathbb{I}_2 \otimes \cdots \otimes \mathbb{I}_N - \mathbb{I}_1 \otimes \Delta_2 \otimes \cdots \otimes \mathbb{I}_N - \ldots - \mathbb{I}_1 \otimes \mathbb{I}_2 \otimes \cdots \otimes \Delta_N$

$\mathbb{V} = -V_1 \otimes \mathbb{I}_2 \otimes \cdots \otimes \mathbb{I}_N - \mathbb{I}_1 \otimes V_2 \otimes \cdots \otimes \mathbb{I}_N - \ldots - \mathbb{I}_1 \otimes \mathbb{I}_2 \otimes \cdots \otimes V_N$

$W = \ldots$ has $\mathcal{O}(N^2)$ terms.

Thus, the nominal separation rank grows as $\mathcal{O}(N^2)$.

It turns out that ...

Separation rank of the Schrödinger operator

Theorem:

The representation of $\mathcal{T} + \mathbb{V}$ to within ϵ in the operator norm has separation rank

$$\mathrm{r}{=}\mathcal{O}\left(\frac{\log(N\,||-\Delta_1 - V_1||/\epsilon)}{\log(1/\mu) - \log(N\,||-\Delta_1 - V_1||/\epsilon)}\right).$$

Let us symmetrically separate $W_{im} = \sum_{k=1}^{r_w} W_i^k\, W_m^k + \mathcal{O}(\epsilon)$. For each value of k the operator has the form $\mathcal{A} = \sum_{i=1}^{N-1} \sum_{m=i+1}^{N} \mathcal{A}_i\, \mathcal{A}_m$.

Theorem:

The representation of $\mathcal{A}$ to within ϵ in the operator norm has separation rank

$$\mathrm{r}{=}\mathcal{O}\left(\frac{\log(N^2\,||\mathcal{A}_1||^2/\epsilon)}{\log(1/\mu) - \log(N^2\,||\mathcal{A}_1||^2/\epsilon)}\right).$$

Thus, the separation rank of the Schrödinger operator grows only as $\log(N)$

Constructive proof

Consider

$$\mathbb{G}(t) = ||-\Delta_1 - V_1|| \bigotimes_{i=1}^{N} (\mathbb{I}_i + t\frac{-\Delta_i - V_i}{||-\Delta_1 - V_1||}),$$

and note that $\mathbb{G}'(0) = \mathcal{T} + \mathbb{V}$. Using finite difference formula of order r, we approximate

$$\mathbb{G}'(0) \approx \sum_{j=1}^{r} \alpha_j \mathbb{G}'(t_j).$$

Similarly, we using $\mathcal{A}_i$ instead of $-\Delta_i - V_i$ in the definition of $\mathbb{G}$, we note that $\mathbb{G}''(0) = 2\mathcal{A}/||\mathcal{A}_1||$ and use the same approach.

(This approximation was first discovered numerically).

Antisymmetry

Since electrons are fermions, the wave function must be antisymmetric, e.g., $\psi(\gamma_2, \gamma_1, \gamma_3, \ldots) = -\psi(\gamma_1, \gamma_2, \gamma_3, \ldots)$, where $\gamma = ((x, y, z), \sigma)$ and σ is the spin.

Given a function of N variables, its "antisymmetrizer" is defined by

$$\mathcal{A} = \frac{1}{N!} \sum_{p \in S_N} (-1)^p \mathcal{P},$$

where S_N is the permutation group on N elements. If $\mathcal{A}$ is applied to a separable function, then the result can be expressed as a Slater determinant,

$$\mathcal{A} \prod_{j=1}^{N} \phi_j(\gamma_j) = \frac{1}{N!} \begin{vmatrix} \phi_1(\gamma_1) & \phi_1(\gamma_2) & \cdots & \phi_1(\gamma_N) \\ \phi_2(\gamma_1) & \phi_2(\gamma_2) & \cdots & \phi_2(\gamma_N) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_N(\gamma_1) & \phi_N(\gamma_2) & \cdots & \phi_N(\gamma_N) \end{vmatrix}.$$

Do we have a problem?

The number of terms in $\mathcal{A}\prod_{j=1}^{N}\phi_j(\gamma_j)$ grows exponentially fast and, although this number can algebraically be reduced somewhat, are we in trouble?

If we care only about computing inner products with $\mathcal{A}\prod_{j=1}^{N}\phi_j(\gamma_j)$, then the so-called Löwdin rules provide a solution,

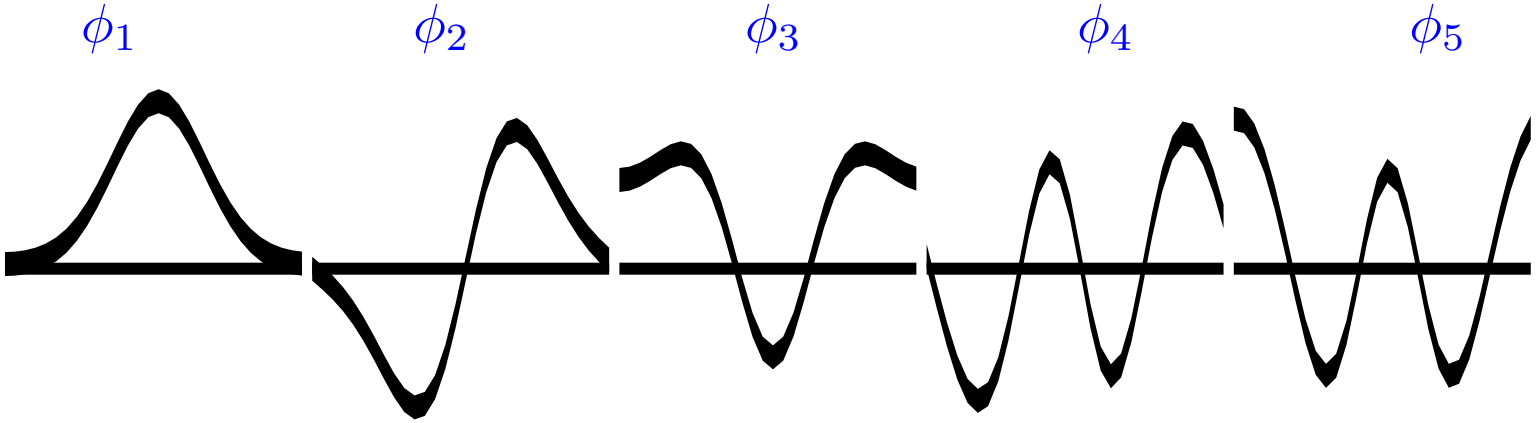
$$\langle\mathcal{A}\prod_{j=1}^{N}\phi_j(\gamma_j),\mathcal{A}\prod_{j=1}^{N}\tilde{\phi}_j(\gamma_j)\rangle=\frac{1}{N!}\begin{vmatrix}\langle\phi_1,\tilde{\phi}_1\rangle & \langle\phi_1,\tilde{\phi}_2\rangle & \cdots & \langle\phi_1,\tilde{\phi}_N\rangle\\ \langle\phi_2,\tilde{\phi}_1\rangle & \langle\phi_2,\tilde{\phi}_2\rangle & \cdots & \langle\phi_2,\tilde{\phi}_N\rangle\\ \vdots & \vdots & \ddots & \vdots\\ \langle\phi_N,\tilde{\phi}_1\rangle & \langle\phi_N,\tilde{\phi}_2\rangle & \cdots & \langle\phi_N,\tilde{\phi}_N\rangle\end{vmatrix}.$$

Computing determinant costs at most $\mathcal{O}(N^3)$, but for large N the matrix is banded and the cost is $\mathcal{O}(N)$, so that we are O.K.

Example of computing the antisymmetric ground state

Separated rank, achieved approximation, and eigenvalue estimates for the separable ($\mathbf{F}_0$) and main approximations to the wavefunction.

	r	ϵ	$\|\|\mathcal{A}\mathbb{H}\mathbf{F}$	$\langle \mathcal{A}\mathbb{H}\mathbf{F}, \mathcal{A}\mathbf{F} \rangle$
$\mathbf{F}_0$	1	$3.4 \cdot 10^{-3}$	322.6727395	322.3859013
$\mathbf{F}$	2	10^{-4}	321.8852595	321.8844158

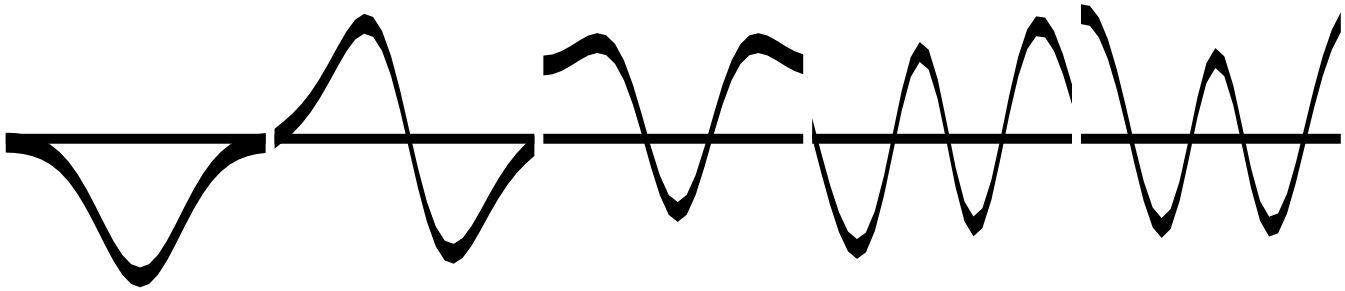


The computed separable approximation $\mathbf{F}_0$ to the wave function.

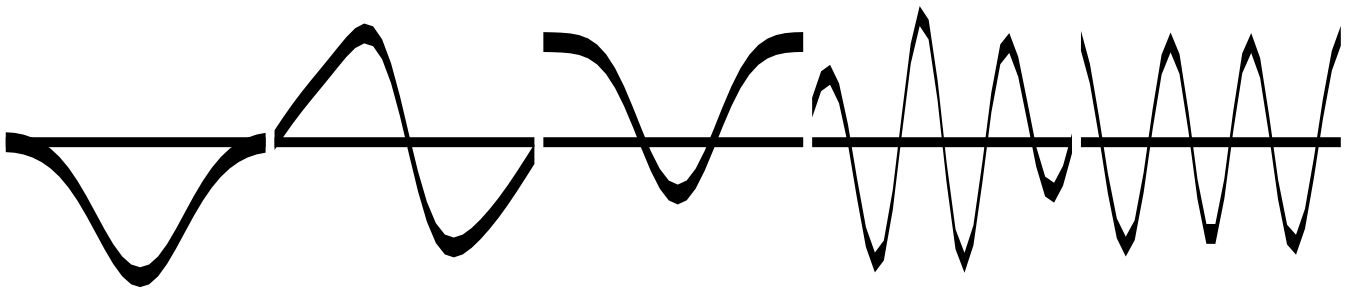
Structure of the antisymmetric ground state

$i = 1$ $i = 2$ $i = 3$ $i = 4$ $i = 5$

$l = 1;\ 0.999350$



$l = 2;\ 0.033093$



Conclusions and future work

- Powerful method for multidimensional problems
- Lattice sums
- Operators in 3D (e.g., oscillatory Green's functions)
- Operators in 6D for multiresolution quantum chemistry (two-electron models)
- We are attempting to solve the multiparticle Schrödinger equation
- Complete MADNESS (Multiresolution ADaptive NumErical Scientific Simulation)