Fast Harmonic Polynomial Transforms

Richard Mikaël Slevinsky

Department of Mathematics,
University of Manitoba

Richard.Slevinsky@umanitoba.ca
https://home.cc.umanitoba.ca/∼slevinrm/

Computational Methods
in Analytics, Inference, and Computation in
Cosmology
Insitut Henri Poincaré

November 29, 2018

Spherical Harmonics

Let μ be a positive Borel measure on $D \subset \mathbb{R}^n$. The inner product:

$$\langle f, g \rangle = \int_D \overline{f(x)} g(x) \, \mathrm{d}\mu(x),$$

induces the norm $\|f\|_2 = \sqrt{\langle f, f \rangle}$ and the Hilbert space $L^2(D, \mathrm{d}\mu(x))$.

Let $\mathbb{S}^2 \subset \mathbb{R}^3$ denote the unit 2-sphere and let $\mathrm{d}\Omega = \sin\theta \, \mathrm{d}\theta \, \mathrm{d}\varphi$.
Then any function $f \in L^2(\mathbb{S}^2, \mathrm{d}\Omega)$ may be expanded in spherical harmonics:

$$f(\theta, \varphi) = \sum_{\ell=0}^{+\infty} \sum_{m=-\ell}^{+\ell} f_\ell^m Y_\ell^m(\theta, \varphi) = \sum_{m=-\infty}^{+\infty} \sum_{\ell=|m|}^{+\infty} f_\ell^m Y_\ell^m(\theta, \varphi),$$

where the expansion coefficients are:

$$f_\ell^m = \frac{\langle Y_\ell^m, f \rangle}{\langle Y_\ell^m, Y_\ell^m \rangle}$$

Spherical Harmonics

For $\ell \in \mathbb{N}_0$ and $|m| \leq \ell$, orthonormal spherical harmonics are defined by:

$$Y_\ell^m(\mathbf{x}) = Y_\ell^m(\theta, \varphi) = \frac{e^{\mathrm{i}m\varphi}}{\sqrt{2\pi}} \underbrace{(-1)^{|m|}\sqrt{(\ell+\tfrac{1}{2})\frac{(\ell-|m|)!}{(\ell+|m|)!}}P_\ell^{|m|}(\cos\theta)}_{\tilde{P}_\ell^{|m|}(\cos\theta)}.$$

Associated Legendre functions are defined by ultraspherical polynomials:

$$P_\ell^m(\cos\theta) = (-2)^m(\tfrac{1}{2})_m \sin^m\theta\, C_{\ell-m}^{(m+\frac{1}{2})}(\cos\theta).$$

The notation $\tilde{P}_\ell^m$ is used to denote orthonormality, and:

$$(x)_n = \frac{\Gamma(x+n)}{\Gamma(x)}$$

is the Pochhammer symbol for the rising factorial.

Spherical Harmonics

Consider the Laplace–Beltrami operator on $\mathbb{S}^2$:

$$\Delta_{\theta,\varphi} = \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}.$$

For $\ell \in \mathbb{N}_0$ and $|m| \leq \ell$, the surface spherical harmonics are the eigenfunctions of $\Delta_{\theta,\varphi}$:

$$\Delta_{\theta,\varphi} Y_\ell^m(\theta,\varphi) = -\ell(\ell+1) Y_\ell^m(\theta,\varphi).$$

Spherical Harmonics

Spherical harmonics satisfy many three-term recurrence relations including:

$$\cos\theta Y_\ell^m = \sqrt{\frac{(\ell-m+1)(\ell+m+1)}{(2\ell+1)(2\ell+3)}}\,Y_{\ell+1}^m + \sqrt{\frac{(\ell-m)(\ell+m)}{(2\ell-1)(2\ell+1)}}\,Y_{\ell-1}^m,$$

$$\sin\theta e^{\mathrm{i}\varphi} Y_\ell^m = \sqrt{\frac{(\ell+m+1)(\ell+m+2)}{(2\ell+1)(2\ell+3)}}\,Y_{\ell+1}^{m+1} - \sqrt{\frac{(\ell-m-1)(\ell-m)}{(2\ell-1)(2\ell+1)}}\,Y_{\ell-1}^{m+1}.$$

They also have an addition theorem:

$$P_\ell(\mathbf{x}\cdot\mathbf{y}) = \frac{4\pi}{2\ell+1}\sum_{m=-\ell}^{+\ell} Y_\ell^m(\mathbf{x})\overline{Y_\ell^m(\mathbf{y})}.$$

Perhaps they have too many properties.

Synthesis and Analysis

A band-limited function of degree-n on the sphere has no nonzero spherical harmonic expansion coefficient of degree-n or greater:

$$f_{n-1}(\theta,\varphi) = \sum_{\ell=0}^{n-1} \sum_{m=-\ell}^{+\ell} f_\ell^m Y_\ell^m(\theta,\varphi).$$

For a spherical harmonic Y_ℓ^m, ℓ is the degree and m is the order.
The transforms of synthesis and analysis convert between representations of a band-limited function in momentum and physical spaces:

Synthesis Sample a band-limited function at a set of points on $\mathbb{S}^2$.

Analysis Convert samples on $\mathbb{S}^2$ to spherical harmonic expansion coefficients.

Normally, an appropriate set of points is chosen to be able to perfectly reconstruct a band-limited function of degree-n.

Synthesis and Analysis

Point sets include:

Equiangular θ_k and φ_j are equispaced and their Cartesian product is taken. These include the celebrated [Driscoll and Healy Jr., 1994] $(2n)(2n-1)$ and [McEwen and Wiaux, 2011] $(n-1)(2n-1)+1$ sampling theorems.

Gaussian $\cos\theta_k$ are Gauss–Legendre points and φ_j are equispaced and their Cartesian product is taken.

HEALPix procedure to generate points that correspond to a hierarchical equal area isolatitude pixelization of $\mathbb{S}^2$ [Górski et al., 2005].

Random with common distributions (e.g. uniformly distributed on $\mathbb{S}^2$).

Data-driven may be treated similar to random.

The naïve cost of synthesis and analysis is $\mathcal{O}(n^4)$ but it can be trivially reorganized to $\mathcal{O}(n^3)$ for isolatitude point sets.

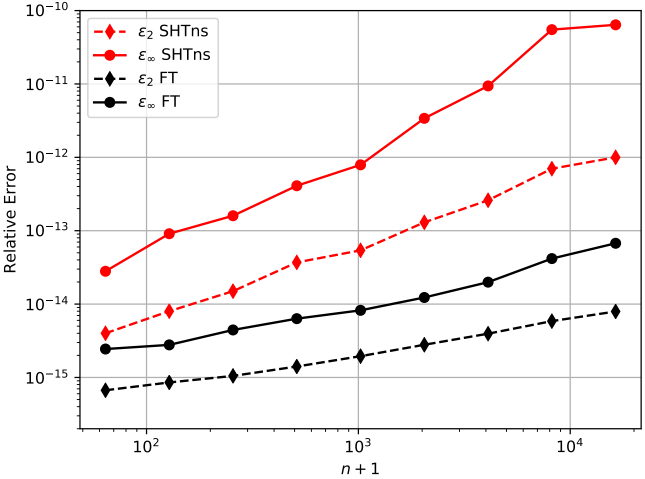
The Connection Problem

University of Manitoba

Complementary to synthesis and analysis is the connection problem: the exact expansion of spherical harmonics in another basis. The sphere is doubly periodic and supports a bivariate Fourier series.

- The (N)FFT solves synthesis and analysis with bivariate Fourier series.
- $Y_\ell^m(\theta,\varphi) = \dfrac{e^{\mathrm{i}m\varphi}}{\sqrt{2\pi}}\tilde{P}_\ell^m(\cos\theta) \Rightarrow$ in longitude, we are done.
- The problem is to convert $\tilde{P}_\ell^m(\cos\theta)$ to Fourier series.
- Since $\tilde{P}_\ell^m(\cos\theta) \propto \sin^m\theta\, C_{\ell-m}^{(m+\frac{1}{2})}(\cos\theta)$, the intuition is that
  - even-ordered $\tilde{P}_\ell^m$ are trigonometric polynomials in $\cos\theta$; and,
  - odd-ordered $\tilde{P}_\ell^m$ are trigonometric polynomials in $\sin\theta$.
- Conversions are *not* one-to-one. The connection problem is an analogue of the McEwen–Wiaux sampling theorem.

Connection vs. Synth. & Anal.



Fast Transforms

Fast transforms should:

- have an $\mathcal{O}(n^2 \log^{\mathcal{O}(1)} n)$ run-time with a similar pre-computation for synthesis and analysis or the connection problem;
- be (backward) stable; and,
- be fast in practice.

Algorithms may be classified:

- as either exact in exact arithmetic and unstable in floating-point arithmetic, and numerically stable algorithms in fixed precision that are approximate to some arbitrarily small tolerance ϵ;
- as either acceleration of synthesis and analysis ($>$10), or acceleration of the connection problem (3); or,
- by analytical apparatus: split-Legendre functions, WKB approximation, the fast multipole method (FMM), and the butterfly algorithm.

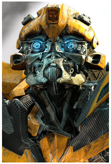
The Transformers


Driscoll


Healy


Mohlenkamp


Suda


Takami


Potts


Steidl


Tasche


Kunis


Rokhlin


Tygert


Slevinsky

Split-Legendre Functions

The Driscoll–Healy $\mathcal{O}(n^2 \log^2 n)$ transform is created by:

- proving the first asymptotically optimal sampling theorem using $(2n)(2n-1)$ equiangular samples. This allows inner products to be represented as discrete sums; if:

$$Z_{k,\ell} = \langle f, T_k P_\ell \rangle,$$

  then they convert $Z_{k,0} = \langle f, T_k \rangle$, obtained by the discrete cosine transform (DCT), to $Z_{0,\ell} = \langle f, P_\ell \rangle$;
- using a technology now known as split-Legendre functions, a factorization of orthogonal polynomial sums evaluated at equiangular points;
- performing a sequence of *masked* and *subsampled* discrete convolutions; and,
- diagonalizing convolutions by the DCT, and fusing different levels in the scheme analytically.

Split-Legendre Functions

Driscoll and Healy also provide a rigorous stability analysis, but the bounds are:

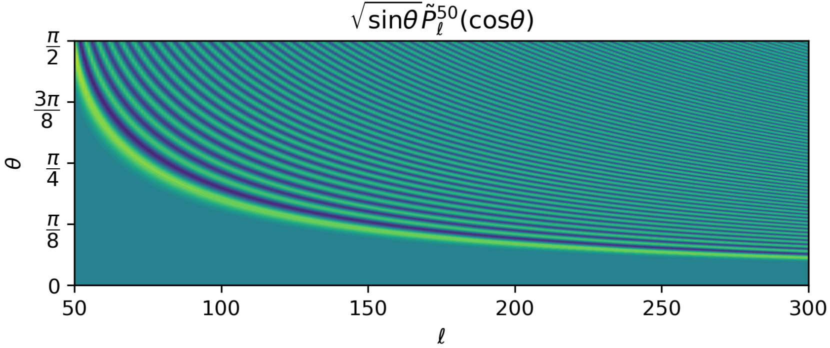
- polynomial in the degree,
- but exponential in the order, or $\mathcal{O}(\ell^m)$,

which explains why the original scheme, even though of foundational importance, is effectively useless.

Algorithms that are exact in exact arithmetic tend to perform poorly in finite precision arithmetic. After the Driscoll–Healy paper, there is a divergence in the literature, where [Potts et al., 1998, Kunis and Potts, 2003, Suda and Takami, 2002, Healy Jr. et al., 2003] proposed *ad hoc* remedies to stabilize the original scheme, and others developed **approximate algorithms that are stable in finite precision arithmetic**. The subsequent algorithms are considered the modern fast transforms.

WKB Asymptotics

As Sturm–Liouville eigenfunctions, it is well-known that associated Legendre functions of high degree and order have a large oscillatory interior as a subset of $\theta \in [0, \pi]$.



WKB Asymptotics

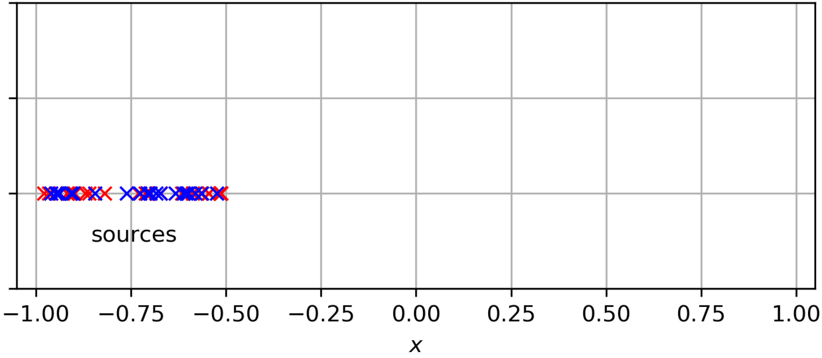
University of Manitoba

In [Mohlenkamp, 1999], a quasi-classical WKB approximation yields:

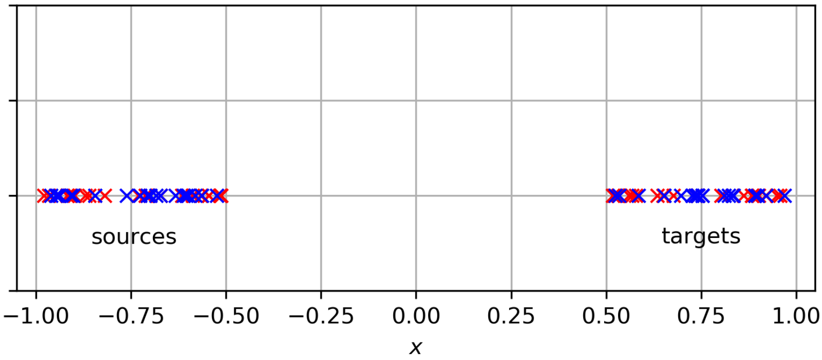
$$\sqrt{\sin\theta}\,\tilde{P}_\ell^m(\cos\theta) \approx \exp\left[\mathrm{i}\int^\theta \sqrt{(\ell+\tfrac{1}{2})^2 - \frac{m^2-\frac{1}{4}}{\sin t}}\,\mathrm{d}t\right].$$

Rigorous improvements can be added to the dominant approximation above, leading to two algorithms for synthesis and analysis with $\mathcal{O}(n^{\frac{5}{2}}\log n)$ and $\mathcal{O}(n^2\log n)$ run-times, respectively. Unfortunately, the numerical evidence does not substantiate the latter compression algorithm.

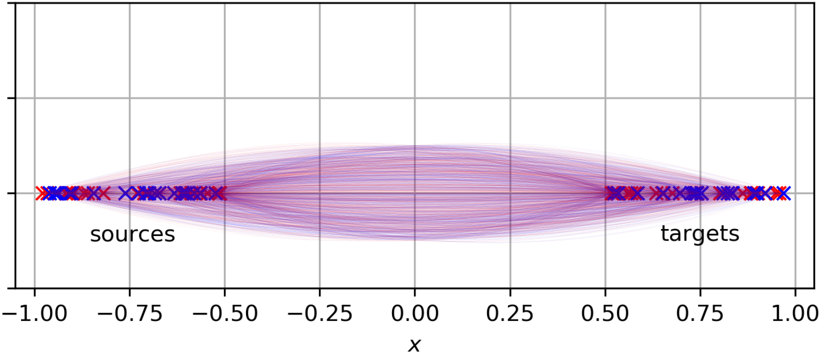
The Fast Multipole Method



The Fast Multipole Method

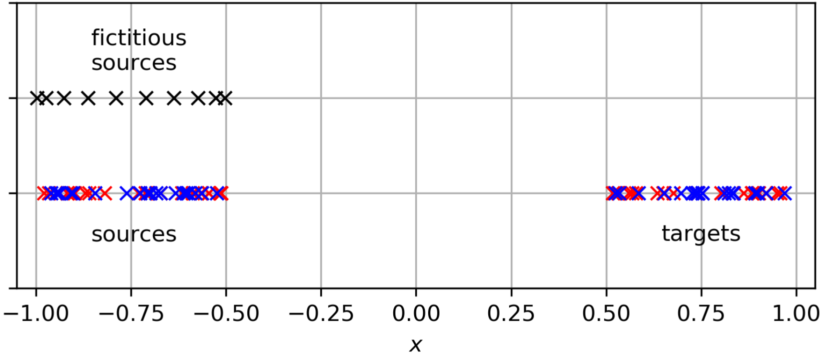


The Fast Multipole Method

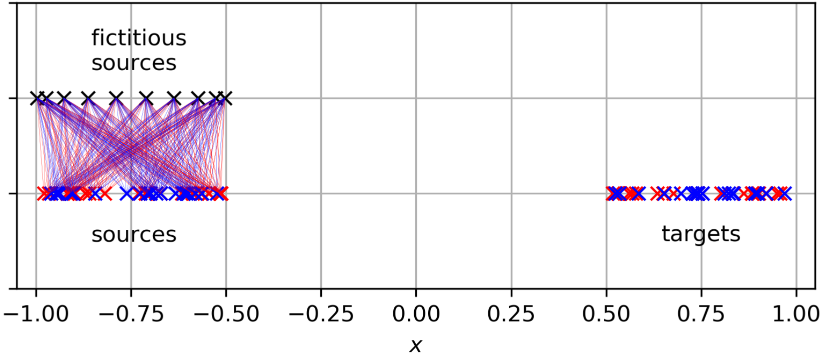


The Fast Multipole Method

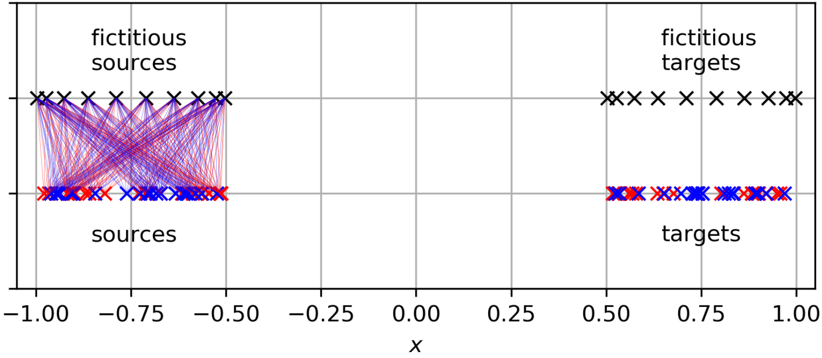
UNIVERSITY OF MANITOBA



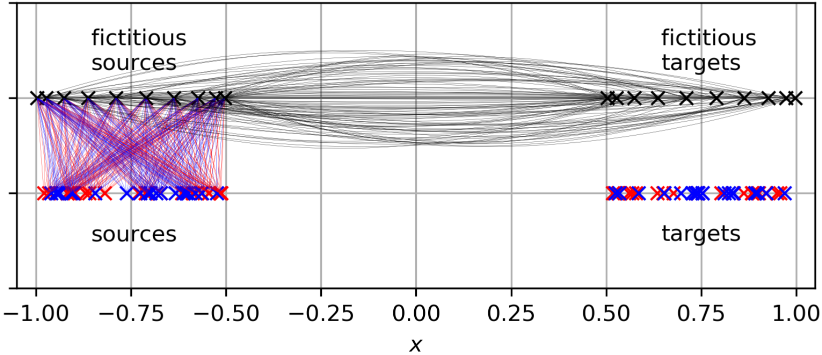
The Fast Multipole Method



The Fast Multipole Method

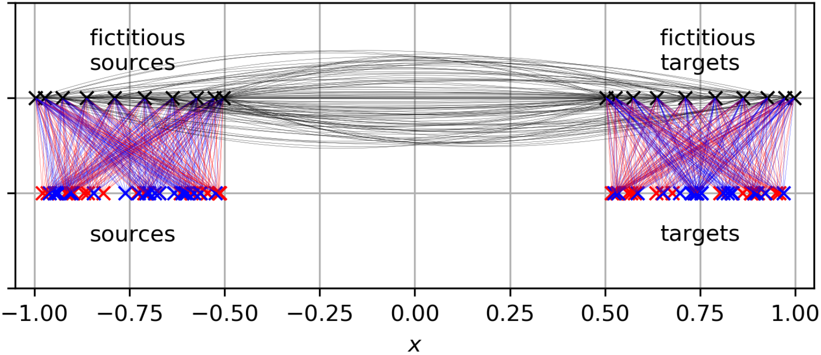


The Fast Multipole Method



The Fast Multipole Method

University of Manitoba



The Fast Multipole Method

University of Manitoba

The numerical method of [Greengard and Rokhlin, 1987] originates from the *multipole expansion* of the Coulombic potential:

$$\frac{1}{|\mathbf{r}-\mathbf{r}_0|} = \frac{1}{\sqrt{r^2 - 2rr_0\cos\theta + r_0^2}} = \frac{1}{r}\sum_{n=0}^{\infty}\left(\frac{r_0}{r}\right)^n P_n(\cos\theta).$$

Expansion for sufficiently small $r_0/r \ll 1$,

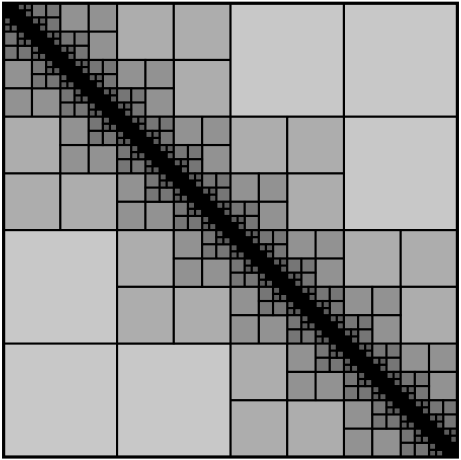
$\Leftrightarrow$ $\mathcal{O}(\log(\epsilon^{-1}))$ terms in the multipole expansion for approximation to precision ϵ,

$\Leftrightarrow$ Subblocks well-separated from the main diagonal.

The FMM enables fast approximate matrix-vector products with Cauchy matrices, generated by sampling $1/(x-y)$ at all pairwise products of $x \in \mathbb{R}^m$ and $y \in \mathbb{R}^n$, and other similar kernels, e.g. [Alpert and Rokhlin, 1991]. What does well-separation resemble?

The Fast Multipole Method

A Cauchy matrix with low-rank subblocks well-separated from the main diagonal:



Eigenfunction Transforms

Associated Legendre functions satisfy:

$$-\frac{\mathrm{d}}{\mathrm{d}x}\left[(1-x^2)\frac{\mathrm{d}}{\mathrm{d}x}\tilde{P}_\ell^m(x)\right]+\frac{m^2}{1-x^2}\tilde{P}_\ell^m(x)=\ell(\ell+1)\tilde{P}_\ell^m(x).$$

The key observation of [Rokhlin and Tygert, 2006] is that the differential equations are structurally similar for $|m|>0$.
For m odd, they expand $\tilde{P}_\ell^m(x)$ in the basis of $\tilde{P}_\ell^1(x)$, and the differential part of the operator reduces to a diagonal scaling, and multiplication by $(1-x^2)$ is a symmetric pentadiagonal matrix with zeros on the sub- and super-diagonal, resulting in a symmetric semiseparable inverse with a chessboard pattern of zeros. Formally:

$$(\mathcal{D}+(m^2-1)\mathcal{M}^{-1})u=\lambda u.$$

Eigenfunction Transforms

University of Manitoba

The Ritz–Galerkin discretization:

$$(\mathcal{D} + (m^2 - 1)\mathcal{M}^{-1})u = \lambda u.$$

Facts:

- The entries of $\mathcal{D}$ are $\ell(\ell+1)$ for $\ell \geq |m|$;
- The entries of

$$[\mathcal{M}^{-1}]_{\ell,n} = \begin{cases} (n+\frac{3}{2})\sqrt{\frac{(\ell+1)(\ell+\frac{3}{2})(\ell+2)}{(n+1)(n+\frac{3}{2})(n+2)}}, & \text{for} \quad \ell \leq n, \quad \ell + n \text{ even}, \\ (\ell+\frac{3}{2})\sqrt{\frac{(n+1)(n+\frac{3}{2})(n+2)}{(\ell+1)(\ell+\frac{3}{2})(\ell+2)}}, & \text{for} \quad \ell > n, \quad \ell + n \text{ even}, \\ 0, & \textit{otherwise}. \end{cases}$$

- The $\tilde{P}_\ell^m$ are the smoothest eigenfunctions and have the smallest eigenvalues.

Diagonal-plus-Semiseparable D & C

UNIVERSITY OF MANITOBA

For full analysis, see [Chandrasekaran and Gu, 2004]. Let $d, u, v \in \mathbb{R}^n$ and:

$$A = D + S = \text{diag}(d) + \text{triu}(uv^\top) + \text{tril}(vu^\top).$$

If:

$$d = \begin{pmatrix} d_1 \\ d_2 \end{pmatrix}, \quad u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \quad \text{and} \quad v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix},$$

then:

$$A = \begin{pmatrix} A_1 & \\ & A_2 \end{pmatrix} + ww^\top,$$

where $w = \begin{pmatrix} u_1 \\ v_2 \end{pmatrix}$, and:

$$A_1 = \text{diag}(d_1) - \text{diag}(u_1)^2 + \text{triu}(u_1(v_1 - u_1)^\top) + \text{tril}((v_1 - u_1)u_1^\top),$$
$$A_2 = \text{diag}(d_2) - \text{diag}(v_2)^2 + \text{triu}(u_2(v_2 - u_2)^\top) + \text{tril}((v_2 - u_2)u_2^\top).$$

Diagonal-plus-Semiseparable D & C

Let $A_1 = Q_1 \Lambda_1 Q_1^\top$ and $A_2 = Q_2 \Lambda_2 Q_2^\top$. Then:

$$A = \begin{pmatrix} A_1 & \\ & A_2 \end{pmatrix} + ww^\top,$$

$$\begin{aligned} A &= \begin{pmatrix} Q_1 \Lambda_1 Q_1^\top & \\ & Q_2 \Lambda_2 Q_2^\top \end{pmatrix} + ww^\top, \\ &= \begin{pmatrix} Q_1 & \\ & Q_2 \end{pmatrix} \left[\begin{pmatrix} \Lambda_1 & \\ & \Lambda_2 \end{pmatrix} + \begin{pmatrix} Q_1 & \\ & Q_2 \end{pmatrix}^\top ww^\top \begin{pmatrix} Q_1 & \\ & Q_2 \end{pmatrix} \right] \begin{pmatrix} Q_1 & \\ & Q_2 \end{pmatrix}^\top, \\ &= \begin{pmatrix} Q_1 & \\ & Q_2 \end{pmatrix} \left[\Delta + zz^\top \right] \begin{pmatrix} Q_1 & \\ & Q_2 \end{pmatrix}^\top, \end{aligned}$$

where $\Delta = \operatorname{diag}(\Lambda_1, \Lambda_2)$ and $z = \begin{pmatrix} Q_1 & \\ & Q_2 \end{pmatrix}^\top w$. The conquer step relates the two subproblems to the larger one via a symmetric diagonal-plus-rank-one eigendecomposition that can be accelerated by the FMM [Gu and Eisenstat, 1995].

Diagonal-plus-Semiseparable D & C

Lemma (Gu and Eisenstat [1995])

Assume that $\delta_1 < \delta_2 < \cdots < \delta_n$ and that $z_j > 0$. Then the eigenvalues $\{\lambda_i\}_{i=1}^n$ of $\Delta + zz^\top$ interlace the diagonal entries:

$$\delta_1 < \lambda_1 < \delta_2 < \lambda_2 < \cdots < \delta_n < \lambda_n,$$

and are the roots of the secular equation:

$$f(\lambda) = 1 + \sum_{j=1}^{n} \frac{z_j^2}{\delta_j - \lambda} = 0.$$

For each eigenvalue λ_i, the corresponding eigenvector is:

$$q_i = \left(\frac{z_1}{\delta_1 - \lambda_i}, \cdots, \frac{z_n}{\delta_n - \lambda_i}\right)^\top \Bigg/ \sqrt{\sum_{j=1}^{n} \frac{z_j^2}{(\delta_j - \lambda_i)^2}}.$$

Eigenfunction Transforms, II

Associated Legendre functions have the symmetric Jacobi matrix:

$$J = \begin{pmatrix} 0 & \beta_1^m & & & \\ \beta_1^m & 0 & \beta_2^m & & \\ & \ddots & \ddots & \ddots & \\ & & \beta_{n-2}^m & 0 & \beta_{n-1}^m \\ & & & \beta_{n-1}^m & 0 \end{pmatrix},$$

where:

$$\beta_\ell^m = \sqrt{\frac{\ell(\ell+2m)}{(2\ell+2m-1)(2\ell+2m+1)}}.$$

If $J = Q\Lambda Q^\top$, the eigenvalues of J are the roots of the $\tilde{P}_{n+m}^m(x)$ and the orthonormal eigenvectors are proportional to the associated Legendre functions evaluate at these roots.

Thus, the eigenvectors Q implement synthesis and their transpose implement analysis at the roots of $\tilde{P}_{n+m}^m(x)$.

Eigenfunction Transforms, II

- But how is synthesis and analysis at a distinct point sets for every order related to a global spherical synthesis and analysis?
- According to [Tygert, 2008], the key to using the Jacobi matrix is post-processing by the Christoffel–Darboux formula, or equivalently the second barycentric formula.
- The discrete eigenfunctions of the Jacobi matrix are $\tilde{P}_{\ell+m}^m(x_k)$, where x_k are the corresponding Gauss–Jacobi quadrature nodes.
- By the barycentric formula and the connection between Gaussian and barycentric weights:

$$p_n(x) = \sum_{k=0}^{n} \frac{\lambda_k f_k}{x - x_k} \Bigg/ \sum_{k=0}^{n} \frac{\lambda_k}{x - x_k},$$

full spherical harmonic synthesis and analysis may be performed at a common Cartesian product point set, accelerated by the FMM.

Tridiagonal D & C

Let the symmetric tridiagonal matrix T be partitioned as:

$$T = \begin{pmatrix} T_1 & a & \\ a^\top & c & b^\top \\ & b & T_2 \end{pmatrix},$$

Then if $T_1 = Q_1 \Lambda_1 Q_1^\top$ and $T_2 = Q_2 \Lambda_2 Q_2^\top$, we have the similarity transformation to a symmetric arrowhead matrix:

$$\begin{pmatrix} Q_1 & & \\ & & 1 \\ & Q_2 & \end{pmatrix}^\top \begin{pmatrix} T_1 & a & \\ a^\top & c & b^\top \\ & b & T_2 \end{pmatrix} \begin{pmatrix} Q_1 & & \\ & & 1 \\ & Q_2 & \end{pmatrix} = \begin{pmatrix} \Lambda_1 & & Q_1^\top a \\ & \Lambda_2 & Q_2^\top b \\ a^\top Q_1 & b^\top Q_2 & c \end{pmatrix}.$$

The symmetric arrowhead spectral decomposition can also be accelerated by the FMM because the eigenvectors are a (normalized) Cauchy matrix of the arrowhead data. This is analyzed by [Gu and Eisenstat, 1994].

Tridiagonal D & C

Lemma (Gu and Eisenstat [1994])

Assume that $a_1 < a_2 < \cdots < a_{n-1}$ *and that* $b_j > 0$. *Then the eigenvalues* $\{\lambda_i\}_{i=1}^n$ *of* $\begin{pmatrix} \mathrm{diag}(a) & b \\ b^\top & c \end{pmatrix}$, *interlace the diagonal entries:*

$$\lambda_1 < a_1 < \lambda_2 < a_2 < \cdots < a_{n-1} < \lambda_n,$$

and are the roots of the secular equation:

$$f(\lambda) = \lambda - c + \sum_{i=1}^{n-1} \frac{b_i^2}{a_i - \lambda} = 0.$$

For each eigenvalue λ_i, *the corresponding eigenvector is:*

$$q_i = \left(\frac{b_1}{\lambda_i - a_1}, \cdots, \frac{b_{n-1}}{\lambda_i - a_{n-1}}, 1 \right)^\top \Bigg/ \sqrt{1 + \sum_{j=1}^{n-1} \frac{b_j^2}{(\lambda_i - a_j)^2}}.$$

Summary

University of Manitoba

Eigenfunction transforms:

- ✓ Accelerate synthesis and analysis or the connection problem to $\mathcal{O}(n^2 \log n)$;
- ✓ Require $\mathcal{O}(n^2 \log n)$ pre-computation, but absurdly large in practice;
- $\sim$ Are stable, but the error is proportional to the 2-norms of the Sturm–Liouville operators, which we know scale as $\mathcal{O}(n^2)$;
- ? Are fast in practice; and,
- $\times$ Have low memory footprint.

The Butterfly Algorithm

University of Manitoba

Unsatisfied with FMM-accelerated eigentransforms, [Tygert, 2010] develops new accelerated synthesis and analysis based on the butterfly algorithm, originating in [Michielssen and Boag, 1996] and studied as an analytical apparatus [O'Neil et al., 2010].

Purpose: Abstract the algebra of the FFT.

Technique: Divide-and-conquer $\Leftrightarrow$ merge-and-split.

Technology: The interpolative decomposition [Liberty et al., 2007].

Proof: Fourier integral operators have *rank-proportional-to-area.*

The ranks of operator compositions are bounded by the smallest rank in the composition, extending applicability beyond Fourier integral operators.

The Interpolative Decomposition

Lemma

Let $A \in \mathbb{R}^{m\times n}$. For any k, there exist $A_{\mathrm{CS}} \in \mathbb{R}^{m\times k}$ whose columns are a unique subset of the columns of A and $A_{\mathrm{I}} \in \mathbb{R}^{k\times n}$ such that:

1. *some subset of the columns of A_{I} makes up the $k \times k$ identity matrix;*
2. $\|\mathrm{vec}(A_{\mathrm{I}})\|_{\infty} \leq 1$;
3. *the spectral norm of A_{I} satisfies $\|A_{\mathrm{I}}\|_2 \leq \sqrt{k(n-k)+1}$;*
4. *the least singular value of A_{I} is at least 1;*
5. *$A_{\mathrm{CS}}A_{\mathrm{I}} = A$ whenever $k = m$ or $k = n$; and,*
6. *when $k < \min\{m, n\}$, the spectral norm of $A - A_{\mathrm{CS}}A_{\mathrm{I}}$ satisfies:*

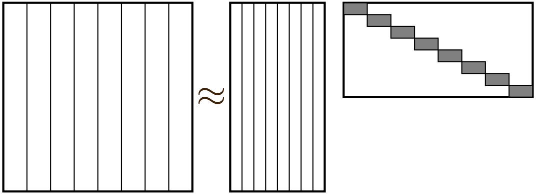
$$\|A - A_{\mathrm{CS}}A_{\mathrm{I}}\|_2 \leq \sqrt{k(n-k)+1}\,\sigma_{k+1}.$$

where σ_{k+1} is the $k+1^{\mathrm{st}}$ singular value of A.

We say that $A \approx A_{\mathrm{CS}}A_{\mathrm{I}}$ and any structure in A is also in A_{CS}.

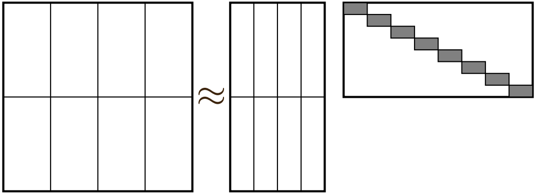
The Butterfly Algorithm

Step 1: Partition $A \in \mathbb{R}^{n \times n}$ into thin strips. Compute IDs of each subblock.



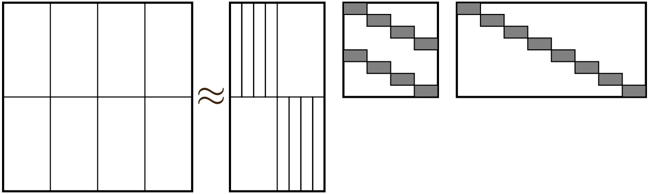
The Butterfly Algorithm

Step 2 (a): Merge the strips and split them approximately in half.



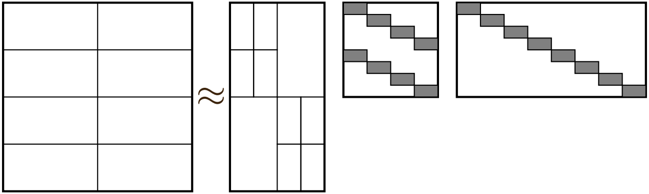
The Butterfly Algorithm

Step 2 (b): Compute IDs of each subblock.



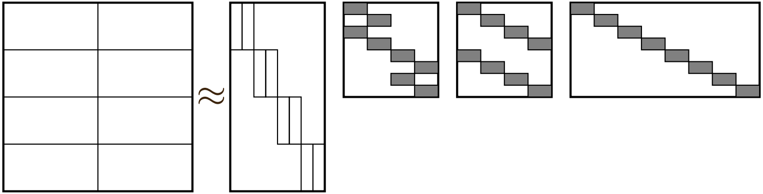
The Butterfly Algorithm

Step 3 (a): Again, merge the strips and split them approximately in half.



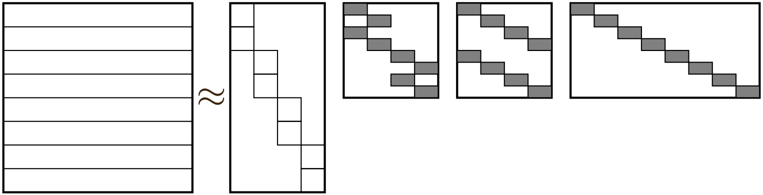
The Butterfly Algorithm

Step 3 (b): Compute IDs of each subblock.



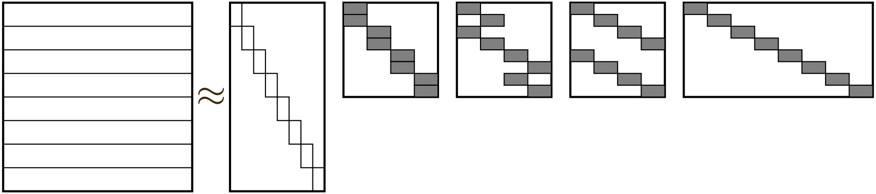
The Butterfly Algorithm

Step 4 (a): Final step, merge the strips and split them approximately in half.



The Butterfly Algorithm

Step 4 (b): Final step, compute IDs of each subblock.



The Butterfly Algorithm

Let $A \in \mathbb{R}^{n \times n}$ have *rank-proportional-to-area*.

- Every time we merge-and-split, the complexity of a matrix-vector product is approximately halved.
- We have created a permuted and sparse block-diagonal factorization.
- Costs $\mathcal{O}(k_{\mathrm{avg}} n^2)$ to compute the factorization and $\mathcal{O}(k_{\mathrm{avg}} n \log n)$ for a matrix-vector product, where k_{avg} is the average rank of all IDs.

To synthesize and analyze associated Legendre functions of all orders requires $\mathcal{O}(k_{\mathrm{avg}} n^3)$ flops to pre-compute and $\mathcal{O}(k_{\mathrm{avg}} n^2 \log n)$ to apply.

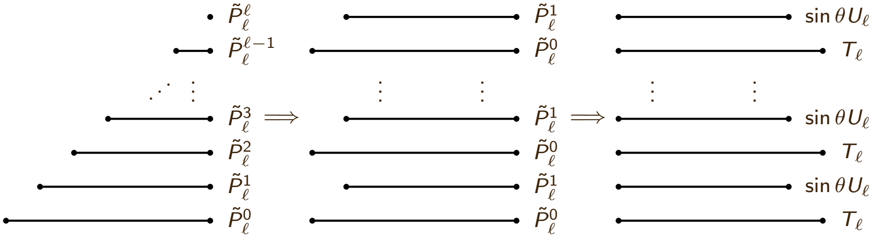
Summary

The butterfly algorithm:

- ✓ Accelerates synthesis and analysis to $\mathcal{O}(n^2 \log n)$;
- × Requires $\mathcal{O}(n^3)$ pre-computation;
- ✓ Is stable, with errors in proportion to the tolerances in the interpolative decompositions;
- ? Is fast in practice; and,
- × Has low memory footprint.

- Tygert's last algorithm is widely used by practitioners including [Seljebotn, 2012] and [Wedi et al., 2013].
- But why does it work?
- And how can we decrease the memory requirements? For $n = 8,192$ Wavemoth requires $212\,\mathrm{GiB}$ to store the butterfly factorizations. For $n \approx 130,000$, the pre-computations are estimated to occupy $45\,\mathrm{TiB}$.

Spherical Harmonics to Fourier



1. Convert high-order layers to layers of order 0 and 1 in $\mathcal{O}(k_{\text{avg}} n^2 \log n)$ flops and $\mathcal{O}(k_{\text{avg}} n^2 \log n)$ storage; and,
2. Convert low-order layers to Fourier series in $\mathcal{O}(n^2 \log n)$ flops and $\mathcal{O}(n \log n)$ storage à la Fast Multipole Method.

The SH Connection Problem

UNIVERSITY OF MANITOBA

Definition

- Let $\{\phi_n(x)\}_{n\geq 0}$ be a family of orthogonal functions with respect to $L^2(\tilde{D}, \mathrm{d}\tilde{\mu}(x))$; and,
- let $\{\psi_n(x)\}_{n\geq 0}$ be another family of orthogonal functions with respect to $L^2(D, \mathrm{d}\mu(x))$.

The connection coefficients:

$$c_{\ell,n} = \frac{\langle \psi_\ell, \phi_n\rangle_{\mathrm{d}\mu}}{\langle \psi_\ell, \psi_\ell\rangle_{\mathrm{d}\mu}},$$

allow for the expansion:

$$\phi_n(x) = \sum_{\ell=0}^{\infty} c_{\ell,n}\psi_\ell(x).$$

The SH Connection Problem

Theorem

Let $\{\phi_n(x)\}_{n\geq 0}$ *and* $\{\psi_n(x)\}_{n\geq 0}$ *be two families of* orthonormal *functions with respect to* $L^2(D, \mathrm{d}\mu(x))$*. Then the connection coefficients satisfy:*

$$\sum_{\ell=0}^{\infty} \overline{c_{\ell,m}} c_{\ell,n} = \delta_{m,n}.$$

- Any matrix $A \in \mathbb{R}^{m\times n}$, $m \geq n$, with orthonormal columns is well-conditioned and Moore–Penrose pseudo-invertible $A^+ = A^\top$.
- For every m, the $\tilde{P}_\ell^m(x)$ are a family of orthonormal functions for the same Hilbert space $L^2([-1,1], \mathrm{d}x)$.

The SH Connection Problem

Definition

Let G_n denote the Givens rotation:

$$G_n = \begin{pmatrix} 1 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & & \vdots \\ 0 & \cdots & c_n & 0 & s_n & \cdots & 0 \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 0 & \cdots & -s_n & 0 & c_n & \cdots & 0 \\ \vdots & & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 1 \end{pmatrix},$$

where the sines and the cosines are in the intersections of the n^{th} and $n+2^{\mathrm{nd}}$ rows and columns, embedded in the identity of a conformable size.

The SH Connection Problem

UNIVERSITY OF MANITOBA

Theorem (Slevinsky [2017a])
The connection coefficients between $\tilde{P}_{n+m+2}^{m+2}(\cos\theta)$ *and* $\tilde{P}_{\ell+m}^{m}(\cos\theta)$ *are:*

$$c_{\ell,n}^m = \begin{cases} (2\ell+2m+1)(2m+2)\sqrt{\dfrac{(\ell+2m)!}{(\ell+m+\frac{1}{2})\ell!}\,\dfrac{(n+m+\frac{5}{2})n!}{(n+2m+4)!}}, & \text{for} \quad \ell \le n, \quad \ell+n \text{ even},\\ -\sqrt{\dfrac{(n+1)(n+2)}{(n+2m+3)(n+2m+4)}}, & \text{for} \quad \ell = n+2,\\ 0, & \text{otherwise}. \end{cases}$$

Furthermore, the matrix of connection coefficients $C^{(m)} = G_0^{(m)} G_1^{(m)} \cdots G_{n-1}^{(m)} G_n^{(m)} I_{(n+3)\times(n+1)}$, *where the sines and cosines are:*

$$s_n^m = \sqrt{\frac{(n+1)(n+2)}{(n+2m+3)(n+2m+4)}}, \quad \text{and} \quad c_n^m = \sqrt{\frac{(2m+2)(2n+2m+5)}{(n+2m+3)(n+2m+4)}}.$$

The SH Connection Problem

"Proof." W.l.o.g., consider $m = 0$ and $n = 5$.

$$C^{(0)} = \begin{pmatrix} 0.91287 & 0.0 & 0.31623 & 0.0 & 0.17593 & 0.0 \\ 0.0 & 0.83666 & 0.0 & 0.39641 & 0.0 & 0.24398 \\ -0.40825 & 0.0 & 0.70711 & 0.0 & 0.3934 & 0.0 \\ 0.0 & -0.54772 & 0.0 & 0.60553 & 0.0 & 0.37268 \\ 0.0 & 0.0 & -0.63246 & 0.0 & 0.5278 & 0.0 \\ 0.0 & 0.0 & 0.0 & -0.69007 & 0.0 & 0.46718 \\ 0.0 & 0.0 & 0.0 & 0.0 & -0.73193 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & -0.76376 \end{pmatrix}$$

The SH Connection Problem

"Proof." Apply the Givens rotation $G_0^{(0)\top}$:

$$G_0^{(0)\top} C^{(0)} = \begin{pmatrix} 1.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.83666 & 0.0 & 0.39641 & 0.0 & 0.24398 \\ 0.0 & 0.0 & 0.7746 & 0.0 & 0.43095 & 0.0 \\ 0.0 & -0.54772 & 0.0 & 0.60553 & 0.0 & 0.37268 \\ 0.0 & 0.0 & -0.63246 & 0.0 & 0.5278 & 0.0 \\ 0.0 & 0.0 & 0.0 & -0.69007 & 0.0 & 0.46718 \\ 0.0 & 0.0 & 0.0 & 0.0 & -0.73193 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & -0.76376 \end{pmatrix}$$

The SH Connection Problem

"Proof." And again:

$$G_1^{(0)\top} G_0^{(0)\top} C^{(0)} = \begin{pmatrix} 1.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 1.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.7746 & 0.0 & 0.43095 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.72375 & 0.0 & 0.44544 \\ 0.0 & 0.0 & -0.63246 & 0.0 & 0.5278 & 0.0 \\ 0.0 & 0.0 & 0.0 & -0.69007 & 0.0 & 0.46718 \\ 0.0 & 0.0 & 0.0 & 0.0 & -0.73193 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & -0.76376 \end{pmatrix}$$

The SH Connection Problem

UNIVERSITY of MANITOBA

"Proof." And again:

$$G_2^{(0)\top} G_1^{(0)\top} G_0^{(0)\top} C^{(0)} = \begin{pmatrix} 1.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 1.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 1.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.72375 & 0.0 & 0.44544 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.68139 & 0.0 \\ 0.0 & 0.0 & 0.0 & -0.69007 & 0.0 & 0.46718 \\ 0.0 & 0.0 & 0.0 & 0.0 & -0.73193 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & -0.76376 \end{pmatrix}$$

The SH Connection Problem

"Proof." And again:

$$G_3^{(0)\top} G_2^{(0)\top} G_1^{(0)\top} G_0^{(0)\top} C^{(0)} = \begin{pmatrix} 1.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 1.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 1.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 1.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.68139 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.6455 \\ 0.0 & 0.0 & 0.0 & 0.0 & -0.73193 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & -0.76376 \end{pmatrix}$$

The SH Connection Problem

"Proof." And again:

$$G_4^{(0)\top} G_3^{(0)\top} G_2^{(0)\top} G_1^{(0)\top} G_0^{(0)\top} C^{(0)} = \begin{pmatrix} 1.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 1.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 1.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 1.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 1.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.6455 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & -0.76376 \end{pmatrix}$$

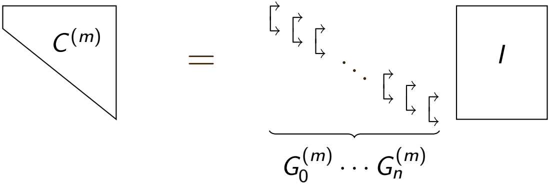
The SH Connection Problem

“Proof.” And finally:

$$G_5^{(0)\top} G_4^{(0)\top} G_3^{(0)\top} G_2^{(0)\top} G_1^{(0)\top} G_0^{(0)\top} C^{(0)} = \begin{pmatrix} 1.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 1.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 1.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 1.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 1.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 1.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \end{pmatrix}$$

The SH Connection Problem

"Proof." Schematically:

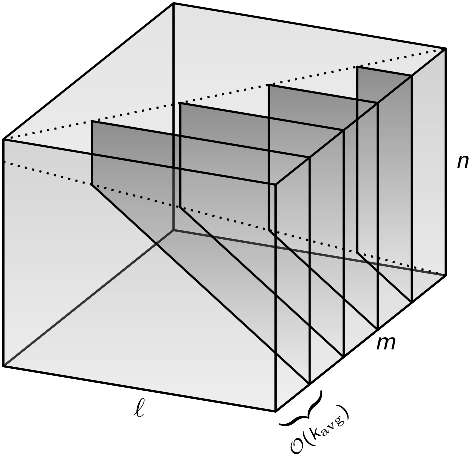


- Conversion between neighbouring layers is $\mathcal{O}(n)$ flops and storage.
- The Givens rotations are computed to high relative accuracy due to analytical expressions of sines and cosines $\Rightarrow$ **backward stable.**

Skeletonizing the Pre-Computation

University of Manitoba

Skeletonizing the pre-computation makes it practical for a laptop.
Neighbouring layers are converted via Given rotations.



Proof That Butter Flies

For $m \in \mathbb{N}$, the connection coefficients between $\tilde{P}_{\ell+2m}^{2m}$ and $\tilde{P}_n^0$ are given by the inner product:

$$c_{\ell,n}^{2m} = \int_{-1}^{1} \tilde{P}_{\ell+2m}^{2m}(x)\tilde{P}_n^0(x)\,\mathrm{d}x.$$

Using the Fourier transform of $\tilde{P}_n^0$:

$$c_{\ell,n}^{2m} = \frac{(-\mathrm{i})^n\sqrt{n+\frac{1}{2}}}{\pi}\int_{\mathbb{R}} j_n(k)\,\mathrm{d}k\int_{-1}^{1} e^{\mathrm{i}kx}\tilde{P}_{\ell+2m}^{2m}(x)\,\mathrm{d}x.$$

The matrix of connection coefficients is an operator composition with the variables $(n \times k) \times (k \times x) \times (x \times \ell)$. The Fourier integral operator is special.

Proof That Butter Flies

UNIVERSITY OF MANITOBA

Theorem (Slevinsky [2017a])

Let:

$$k_1(n,\varepsilon) := 2\left(\frac{\varepsilon}{2}\sqrt{\frac{\pi}{2n+1}}(n+1)\Gamma(n+\tfrac{3}{2})\right)^{\frac{1}{n+1}},$$

and let:

$$k_2(\ell,m,n,\varepsilon) := \frac{1}{8}\left(\frac{2}{\varepsilon}\sqrt{\frac{2n+1}{\pi}}\sqrt{\frac{(2\ell+4m+1)\Gamma(\ell+4m+1)}{\Gamma(\ell+1)}}\frac{1}{m\Gamma(m+\frac{1}{2})}\right)^{\frac{1}{m}}.$$

Then only integration over $k_1(n,\varepsilon) \leq |k| \leq k_2(\ell,m,n,\varepsilon)$ contributes to:

$$c_{\ell,n}^{2m} = \frac{(-\mathrm{i})^n\sqrt{n+\frac{1}{2}}}{\pi}\int_{\mathbb{R}} j_n(k)\,\mathrm{d}k\int_{-1}^{1} e^{\mathrm{i}kx}\tilde{P}_{\ell+2m}^{2m}(x)\,\mathrm{d}x.$$

to precision $\varepsilon > 0$.

Summary

University of Manitoba

The butterfly algorithm applied to the connection problem:

- ✓ Requires an $\mathcal{O}(n^2 \log^2 n)$ run-time;
- × Requires $\mathcal{O}(n^3 \log n)$ pre-computation but only $10\times$ more expensive than run-time in practice;
- ✓ Is (backward) stable, with error scaling as $\mathcal{O}(\sqrt{n}\epsilon)$ for the slow transform and $\mathcal{O}(n\epsilon)$ for the fast transform;
- ? Is fast in practice; and,
- ✓ Has low memory footprint.

Eigenfunction Transforms, III

Associated Legendre functions also satisfy:

$$-(1-x^2)\frac{\mathrm{d}}{\mathrm{d}x}\left[(1-x^2)\frac{\mathrm{d}}{\mathrm{d}x}\tilde{P}_\ell^m(x)\right] + m^2\tilde{P}_\ell^m(x) = \ell(\ell+1)(1-x^2)\tilde{P}_\ell^m(x).$$

Expanding $\tilde{P}_\ell^\mu(x)$ in the basis of $\tilde{P}_\ell^m(x)$ (for $\mu - m$ even), we find formally:

$$(\mathcal{M}\mathcal{D} + (\mu^2 - m^2)\mathcal{I})u = \lambda\mathcal{M}u.$$

For large μ and m, the entries of $\mathcal{M}^{-1}$ are still semiseparable but prone to severe (factorial) scaling. This rules out the diagonal-plus-semiseparable eigensolvers, but how can we reintroduce symmetry?
The key observation of [Slevinsky, 2017b] is that multiplication by $1-x^2$ is a symmetric positive-definite operator and thus has a Cholesky factorization $\mathcal{M} = \mathcal{R}^\top\mathcal{R}$. Letting $u = \mathcal{R}^\top v$ and multiplying from the left by $\mathcal{R}^{-\top}$, we arrive at:

$$(\mathcal{R}\mathcal{D}\mathcal{R}^\top + (\mu^2 - m^2)\mathcal{I})v = \lambda\mathcal{R}\mathcal{R}^\top v.$$

Eigenfunction Transforms, III

Facts about $(\mathcal{R}\mathcal{D}\mathcal{R}^\top + (\mu^2 - m^2)\mathcal{I})v = \lambda\mathcal{R}\mathcal{R}^\top v$:

- The entries of $\mathcal{D}$ are $\ell(\ell+1)$ for $\ell \geq |m|$;
- The Cholesky factor $\mathcal{R}$ is:

$$\mathcal{R} = \begin{pmatrix} c_1^m & 0 & d_1^m & & \\ & \ddots & \ddots & \ddots & \\ & & c_{n-2}^m & 0 & d_{n-2}^m \\ & & & c_{n-1}^m & 0 \\ & & & & c_n^m \end{pmatrix},$$

where:

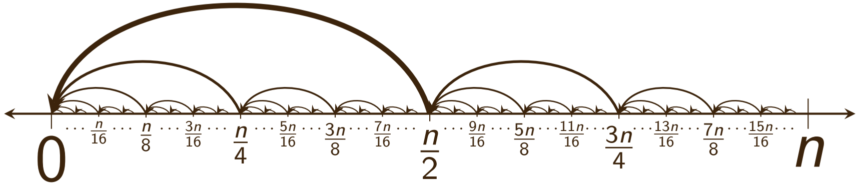
$$c_\ell^m = \sqrt{\frac{(\ell+2m)(\ell+2m+1)}{(2\ell+2m-1)(2\ell+2m+1)}}, \quad \text{and} \quad d_\ell^m = -\sqrt{\frac{\ell(\ell+1)}{(2\ell+2m+1)(2\ell+2m+3)}}.$$

- Symmetric-definite tridiagonal D & C is proposed by [Borges and Gragg, 1993].

Eigenfunction Transforms, III

This symmetric-definite banded generalized eigenvalue problem:

- allows for use of arrowhead divide-and-conquer algorithms to allow $\mathcal{O}(n \log n)$ application of the connection problem; and,
- allows for the pre-computation to be recursively subdivided and reduced from a total cost of $\mathcal{O}(n^2 \log n)$ down to $\mathcal{O}(n^{\frac{3}{2}} \log n)$, which is superoptimal, based on $\mathcal{O}(\log \sqrt{n})$ levels in the following schematic.



The Story So Far

University of Manitoba

- A great thrust has been made to create asymptotically fast spherical harmonic transforms that are practical. Are they faster than the fastest slow methods of [Schaeffer, 2013, Reinecke and Seljebotn, 2013, Ishioka, 2018]? Not yet.
- All technologies are all similar in spirit: they divide and conquer in the presence of oscillations on harmonic polynomials that are separable Sturm–Liouville eigenfunctions of Δ with well-separated spectra.
- Is there a backward stable method? Yes [Slevinsky, 2017a].
- Is there a pre-computation-free method? Yes [Slevinsky, 2017b].
- Is there free and open source software? Yes, including `FastTransforms.jl` in Julia (experimental quality) and `FastTransforms` in C (production quality).
- Many of the technological apparatuses have analogues for other 2D harmonic polynomials: on the disk [Zernike, 1934], triangle [Proriol, 1957], rectangle, deltoid, wedge, etc. . . , as well as spin-weighted spherical harmonics, allowing for a unified framework.

References (1)

J. R. Driscoll and D. M. Healy Jr. Computing Fourier transforms and convolutions on the 2-sphere.
Adv. Appl. Math., 15:202–250, 1994.

J. D. McEwen and Y. Wiaux.
A novel sampling theorem on the sphere.
IEEE Trans. Sig. Proc., 59:5876–5887, 2011.

K. M. Górski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelmann.
HEALPix: A framework for high-resolution discretization and fast analysis of data distributed on the sphere.
Ap. J., 622:759–771, 2005.

D. Potts, G. Steidl, and M. Tasche.
Fast and stable algorithms for discrete spherical Fourier transforms.
Linear Algebra Appl., 275–276:433–450, 1998.

References (2)

S. Kunis and D. Potts.
Fast spherical Fourier algorithms.
J. Comp. Appl. Math., 161:75–98, 2003.

R. Suda and M. Takami.
A fast spherical harmonics transform algorithm.
Math. Comp., 71:703–715, 2002.

D. M. Healy Jr., D. N. Rockmore, P. J. Kostelec, and S. Moore.
FFTs for the 2-sphere–improvements and variations.
J. Fourier Anal. Appl., 9:341–385, 2003.

M. J. Mohlenkamp.
A fast transform for spherical harmonics.
J. Fourier Anal. Appl., 5:159–184, 1999.

L. Greengard and V. Rokhlin.
A fast algorithm for particle simulations.
J. Comp. Phys., 73:325–348, 1987.

References (3)

UNIVERSITY OF MANITOBA

B. K. Alpert and V. Rokhlin.
A fast algorithm for the evaluation of Legendre expansions.
SIAM J. Sci. Stat. Comput., 12:158–179, 1991.

V. Rokhlin and M. Tygert.
Fast algorithms for spherical harmonic expansions.
SIAM J. Sci. Comput., 27:1903–1928, 2006.

S. Chandrasekaran and M. Gu.
A divide-and-conquer algorithm for the eigendecomposition of symmetric block-diagonal plus semiseparable matrices.
Numer. Math., 96:723–731, 2004.

M. Gu and S. C. Eisenstat.
A divide-and-conquer algorithm for the symmetric tridiagonal eigenproblem.
SIAM J. Matrix Anal. Appl., 16:172–191, 1995.

M. Tygert.
Fast algorithms for spherical harmonic expansions, II.
J. Comp. Phys., 227:4260–4279, 2008.

References (4)

M. Gu and S. C. Eisenstat.
A stable and efficient algorithm for the rank-one modification of the symmetric eigenproblem.
SIAM J. Matrix Anal. Appl., 15:1266–1276, 1994.

M. Tygert.
Fast algorithms for spherical harmonic expansions, III.
J. Comp. Phys., 229:6181–6192, 2010.

E. Michielssen and A. Boag.
A multilevel matrix decomposition algorithm for analyzing scattering from large structures.
IEEE Trans. Antennas Propagat., 44:1086–1093, 1996.

M. O'Neil, F. Woolfe, and V. Rokhlin.
An algorithm for the rapid evaluation of special function transforms.
Appl. Comput. Harmon. Anal., 28:203–226, 2010.

References (5)

University of Manitoba

E. Liberty, F. Woolfe, P.-G. Martinsson, V. Rokhlin, and M. Tygert.
Randomized algorithms for the low-rank approximation of matrices.
Proc. Nat. Acad. Sci., 104:20167–20172, 2007.

D. S. Seljebotn.
Wavemoth–fast spherical harmonic transforms by butterfly matrix compression.
Astro. J. Suppl. Series, 199:12, 2012.

N. P. Wedi, M. Hamrud, and G. Mozdzynski.
A fast spherical harmonics transform for global NWP and climate models.
Monthly Weather Review, 141:3450–3461, 2013.

R. M. Slevinsky.
Fast and backward stable transforms between spherical harmonic expansions and bivariate Fourier series.
Appl. Comput. Harmon. Anal., 2017a.

References (6)

R. M. Slevinsky.
Conquering the pre-computation in two-dimensional harmonic polynomial transforms.
arXiv:1711.07866, 2017b.

C. F. Borges and W. B. Gragg.
A parallel divide and conquer algorithm for the generalized real symmetric definite tridiagonal eigenproblem.
In L. Reichel, A. Ruttan, and R. S. Varga, editors, *Numerical Linear Algebra and Scientific Computing*, pages 11–29. de Gruyter, 1993.

N. Schaeffer.
Efficient spherical harmonic transforms aimed at pseudospectral numerical simulations.
Geochem. Geophys. Geosyst., 14:751–758, 2013.

M. Reinecke and D. S. Seljebotn.
Libsharp – spherical harmonic transforms revisited.
arXiv:1303.4945, 2013.

References (7)

K. Ishioka.
A new recurrence formula for efficient computation of spherical harmonic transform.
J. Met. Soc. Japan, 96:241–249, 2018.

F. Zernike.
Beugungstheorie des Schneidenverfahrens und seiner verbesserten Form, der Phasenkontrastmethode.
Physica, 1:689–704, 1934.

J. Proriol.
Sur une famille de polynômes à deux variables orthogonaux dans un triangle.
C. R. Acad. Sci. Paris, 245:2459–2461, 1957.