Numerical analysis and orthogonal polynomials

CHAPTER 2: Orthogonal expansions

Arieh Iserles

OPTRIM, Madrid

October 2017
List of topics:

2.1 A primer of spectral methods

2.2 Fast Legendre Transform via hypergeometric identities

2.3 Connection coefficients

2.4 Fast Legendre Transform via perturbed Chebyshev grid

2.5 Fast Multipole Method (FMM)

2.6 FMM applied to fast transforms
2.1 A primer of spectral methods

Classical methods for differential equations:

**Finite differences:** Discretise your PDE on a grid, approximating derivatives by finite linear combinations of nodal values.

**Finite elements:** Approximate the solution as a linear combination of functions with small support and *either* minimise a variational functional (*the Ritz method*) or determine a projection of the weak solution (*the Galerkin method*).

The end-product in each case is a large system (linear or nonlinear) of *sparse* algebraic equations: solving this system represents the real numerical price tag of a method.

**An alternative paradigm:**

Expand the solution in an orthogonal system so that convergence occurs very rapidly. Although the outcome is a *dense* algebraic system, typically it is considerably smaller.
We seek an orthogonal basis \( \{ \phi_n \} \) of the ambient space s.t.*

1. The expansion coefficients

\[
\hat{f}_m = \frac{\int_a^b f(x) \phi_m(x) \, dx}{\int_a^b \phi_m^2(x) \, dx}
\]

converge rapidly;

2. The first \( n \) coefficients \( \hat{f}_m \) can be evaluated rapidly: this typically means in \( \mathcal{O}(n(\log n)^q) \) operations for some small \( q \).

3. Let \( \hat{f} \in \mathbb{C} \) be the vector of the coefficients. We say that \( \mathcal{D} \) is a differentiation matrix if \( \hat{f}' = \mathcal{D}\hat{f} \). It need be known explicitly, it should be cheap to form \( \mathcal{D}v \) for a vector \( v \) and, in an ideal world, we want \( \mathcal{D} \) to be skew-Hermitian.

*We assume for simplicity that the ambient space is \( L(a, b) \).
4. The basis functions \( \{ \phi_n \} \) match required boundary conditions (e.g. Dirichlet, Neumann or periodic).

**Example I** The diffusion equation \( u_t = u_{xx} \) for \(-1 \leq x \leq 1, \ t \geq 0\), with periodic boundary conditions.

We choose the **Fourier basis** \( \phi_n = e^{i\pi nx}, \ n \in \mathbb{Z} \) – in other words, represent

\[
    u(x, t) \approx \sum_{n=-N+1}^{N} \hat{u}_n(t)e^{i\pi nx}
\]

for some \( N \gg 1 \). Trivially, we end up with \( \hat{u}'_n = -\pi^2 n^2 \hat{u}_n \), in other words with

\[
    u(x, t) \approx \sum_{m=-N+1}^{N} \alpha_m e^{i\pi nx - \pi^2 n^2 t},
\]

where the \( \alpha_m \)s are given by the Fourier expansion of the initial condition – this can be done in \( \mathcal{O}(N \log N) \) operations by **Fast Fourier Transform (FFT)**.
In this case the spectral method is nothing else but separation of variables, although in general this need not be true.

Let us examine points 1–4:

1. The $\alpha_n$s converge for a smooth function at a spectral speed (i.e., their size is asymptotically smaller than a reciprocal of any polynomial in $|n|$). This means that we need a fairly small $N$.

2. $\alpha_n$ for $-N + 1 \leq n \leq N$ can be calculated in $O(N \log N)$ flops by FFT.

3. The differentiation matrix $D$ is diagonal and skew-Hermitian, $D'_{n,n} = -i\pi n D_{n,n}$.

4. Periodicity is matched by the basis functions.

Example II The diffusion equation $u_t = u_{xx}$ for $-1 \leq x \leq 1$, $t \geq 0$, with Dirichlet boundary conditions.

The obvious choice is a Chebyshev basis $\{T_n\}_{n=0}^{\infty}$, whereby

$$\hat{u}_0 = \frac{1}{\pi} \int_0^{\pi} f(\cos \theta) \, d\theta, \quad \hat{u}_n = \frac{2}{\pi} \int_0^{\pi} f(\cos \theta) \cos n\theta \, d\theta, \quad n \in \mathbb{N}.$$
\[ T'_{2n} = 4n \sum_{\ell=0}^{n-1} T_{2\ell+1}, \quad T'_{2n+1} = (2n + 1)T_0 + 2(2n + 1) \sum_{\ell=1}^{n} T_{2\ell} \]

implies

\[ T''_{2n} = 4n^3T_0 + 8n \sum_{k=1}^{n-1} (n^2 - k^2)T_{2k}, \]

\[ T''_{2n+1} = 4(2n + 1) \sum_{k=0}^{n-1} (n - k)(n + k + 1)T_{2k+1}. \]

Consequently, letting \( u(x, t) \approx \sum_{n=0}^{N} \hat{u}_n(t)T_n(x) \), we have

\[ \hat{u}'_0 = 4 \sum_{k=0}^{[N/2]} k^3 \hat{u}_{2k}, \]

\[ \hat{u}'_{2n} = 8 \sum_{k=n+1}^{[N/2]} k(k^2 - n^2)\hat{u}_{2k}, \quad n \geq 1, \]

\[ \hat{u}'_{2n+1} = 4 \sum_{k=n+1}^{[N/2]-1} (2k + 1)(k - n)(k + n + 1)\hat{u}_{2k+1}. \]
A checklist:

1. An expansion of a smooth function in Chebyshev polynomials converges at spectral speed.
2. Chebyshev expansion of initial condition can be computed with Fast Cosine Transform in $O(N \log N)$ operations.
3. The differentiation matrix of any polynomial basis is necessarily lower-triangular. A naive product by a vector is expensive and, worse, lower triangular is as far from skew-Hermitian as they come – and this is bad news for numerical stability – we might be compelled to solve the ODE using excessively small time steps!
4. We can match Dirichlet (or Neumann, Robin, . . . ) boundary conditions with Chebyshev polynomials.

Point 3 is a serious impediment toward use of Chebyshev spectral methods for parabolic PDEs (although there are clever ways to overcome it), but neither for elliptic not hyperbolic PDEs.
Example III Consider the spectral problem for a Fredholm operator,

\[ \int_{-1}^{1} f(x) K(y-x) \, dx = \lambda f(y), \quad -1 \leq y \leq 1. \]

This is a compact operator hence it has a point spectrum with at most a single accumulation point at the origin. The finite section method approximates

\[ f(x) \approx \sum_{n=0}^{\infty} \alpha_n \phi_n(x), \]

where \( \{\phi_n\}_{n=0}^{\infty} \) is an orthonormal basis of \( L[-1, 1] \). It is easy to prove that

\[ \sum_{n=0}^{\infty} \alpha_n \int_{-1}^{1} \int_{-1}^{1} \phi_n(x) \phi_m(y) K(y-x) \, dx \, dy = \lambda \alpha_m, \quad m \in \mathbb{Z}_+, \]

i.e. the algebraic eigenvalue problem \( A[\infty] \alpha = \lambda \alpha \), where

\[ A_{m,n}^{[\infty]} = \int_{-1}^{1} \int_{-1}^{1} \phi_n(x) \phi_m(y) K(y-x) \, dx \, dy, \quad m, n \in \mathbb{Z}_+. \]
Denoting by $A^{[N]}$ the $(N + 1) \times (N + 1)$ section of $A^{[\infty]}$ (i.e., restricting the range of $m, n$ to $0, \ldots, n$) we obtain the finite-dimensional eigenvalue problem $A^{[N]} \alpha^{[N]} = \lambda \alpha^{[N]}$, which can be solved by standard software. Clearly, the faster the convergence, the smaller we need to take $N$, the smaller the eigenvalue problem.

Which basis to choose? In this case the nature of the basis is dictated by the problem itself: we need to expand in Legendre polynomials.

1. The bivariate expansion converges at a spectral speed.
2. Just now we have no fast way of computing a Fast Legendre Transform but watch this space!
3. A differentiation matrix is irrelevant in this setting.
4. Using Legendre polynomials, we have no problems with boundary conditions.
Speed of convergence

We will examine convergence for two functions: $e^{-x^2}$ and $e^{\cos \pi x}$, both in $[-1, 1]$: note that the second is periodic.

$e^{-x^2}$:

\[ \log_{10} \left| \sum_{|n| \leq N} \hat{f}_n \psi_n - f \right| \]

\[ \log_{10} |\hat{f}_n| \]
$e^{\cos \pi x}$:

**Fourier**

**Chebyshev**

$\log_{10} \left| \sum_{|n| \leq N} \hat{f}_n \psi_n - f \right|$

$\log_{10} |\hat{f}_n|$

$N = 16$

$N = 32$

$N = 64$

$N = 128$

$N = 256$
**Comments:**

A. Once $e^{-x^2}$, which is not periodic, is approximated by Fourier, the error decays like $O\left(|N|^{-1}\right)$: very slowly. Otherwise, in all instances the error decays spectrally.

B. Once the error decays spectrally, we need a fairly small number of terms – the outcome is a very small algebraic system.

Our presentation is exclusively for linear problems but spectral methods can deal with nonlinear equations, e.g. by exploiting the identities

$$f(x)g(x) = \sum_{n=-\infty}^{\infty}\left(\sum_{m=-\infty}^{\infty}\hat{f}_m\hat{g}_{n-m}\right)e^{inx} \quad \text{(convolution)},$$

$$T_m T_n = \frac{1}{2}(T_{|m-n|} + T_{m+n}).$$

They can also be easily generalised by tensor products to parallelepipeds.*

*More exotic domains in $\mathbb{R}^m, m \geq 2$, are much more complicated.
2.2 Fast Legendre Transform via hypergeometric identities

We can do Chebyshev expansions (of the first and the second kind) fast – what about Legendre polynomials (and, with greater generality, Jacobi polynomials)? There are several algorithms and the first is based on a very strange hypergeometric identity.

Step I: \( \frac{x^n}{n!} = \frac{1}{2^n} \sum_{m=0}^{\lfloor n/2 \rfloor} \frac{2n + 1 - 4m}{m!(\frac{3}{2})_{n-m}} P_{n-2m}(x). \)

Step II: Let \( f \) be analytic within the Bernstein ellipse \( B_r = \left\{ \frac{1}{2}(re^{i\theta} + r^{-1}e^{-i\theta}) : -\pi \leq \theta \leq \pi \right\}, \quad r \in (0, 1). \)

Then \( \hat{f}_n = (n + \frac{1}{2}) \sum_{m=0}^{\infty} \frac{f(n+2m)(0)}{2^{n+2m}m!(\frac{1}{2})_{n+m+1}}, \quad n \in \mathbb{Z}_+. \)
Step III: If \( \gamma \) is a simple, closed, positively oriented Jordan curve in \( B_r \setminus [-1, 1] \), then by the Cauchy integral theorem

\[
\hat{f}_n = \frac{c_n}{2\pi i} \oint_{\gamma} \frac{f(z)}{z^{n+1}} \varphi_n(z) \, dz, \quad n \in \mathbb{Z}_+,
\]

where

\[
c_n = \frac{n!}{2^n (\frac{1}{2})^n}, \quad \varphi_n(z) = 2 \, _2F_1 \left[ \begin{array}{c} \frac{n+1}{2}, \frac{n+2}{2} \\ n + \frac{3}{2} \end{array} ; \frac{1}{z^2} \right].
\]

Step IV: \( \varphi_n \) is a slowly-convergent Taylor series. Instead use the nonlinear transformation

\[
_2F_1 \left[ \begin{array}{c} a, a + \frac{1}{2} \\ c; \end{array} ; 2\zeta - \zeta^2 \right] = (1 - \frac{1}{2}\zeta)^{-2a} \, _2F_1 \left[ \begin{array}{c} 2a, 2a - c + 1 \\ c; \end{array} ; \frac{\zeta}{2 - \zeta} \right],
\]

where \( \text{Re} \zeta < 1 \).
Set \( a = \frac{1}{2}(n + 1) \), \( c = n + \frac{3}{2} \), whereby

\[
\varphi_n((2\zeta - \zeta^2)^{-1/2}) = \frac{1}{(1 - \frac{1}{2}\zeta)^n+1} {\binom{n+\frac{1}{2}}{n+\frac{3}{2}}} \left[ \frac{\zeta}{2 - \zeta} \right].
\]

Choose \( \gamma \) as \( B_\rho, \rho \in (r, 1) \) (therefore \( B_\rho \) lives inside \( B_r \)). Then

\[ \left| \frac{\zeta}{2 - \zeta} \right| = \rho^2 < 1 \]

and the hypergeometric function decays rapidly.

**Step V:** After some algebra

\[
\hat{f}_n = \frac{(2\rho)^nc_n}{2\pi} \int_{-\pi}^{\pi} (1 - \rho^2 e^{2i\theta}) f\left(\frac{1}{2}(\rho e^{i\theta} + \rho^{-1} e^{-i\theta})\right) e^{in\theta} \chi_n(\rho^2 e^{2i\theta}) d\theta,
\]

where

\[
\chi_n(z) = {\binom{n+\frac{1}{2}}{n+\frac{3}{2}}} z.
\]
**Step VI:** The Taylor coefficients \( \chi_{n,m} = (n + 1)_m (\frac{1}{2})_m/[m!(n + \frac{3}{2})_m] \) converge rapidly and the first \( M \) terms can be evaluated by recursion in \( \mathcal{O}(M) \) operations. Truncating the expansion

\[
\hat{f}_n \approx \frac{(2\rho)^n c_n}{2\pi} \int_{-\pi}^{\pi} (1 - \rho^2 e^{2i\theta}) f\left(\frac{\rho e^{i\theta} + \rho^{-1} e^{-i\theta}}{2}\right) e^{in\theta} \sum_{m=0}^{M} \chi_{n,m} \rho^{2m} e^{2im\theta} d\theta
\]

\[
= (2\rho)^n c_n \sum_{m=0}^{M} \chi_{n,m} \rho^{2m} \hat{v}_{n+2m} := \hat{f}_n^{[M]},
\]

where

\[
\hat{v}_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} (1 - \rho^2 e^{2i\theta}) f\left(\frac{\rho e^{i\theta} + \rho^{-1} e^{-i\theta}}{2}\right) e^{in\theta} d\theta, \quad -N + 1 \leq m \leq N,
\]

can be computed by **FFT** in \( \mathcal{O}(N \log N) \) operations.
Convergence

To compute the Taylor coefficients

\[ \chi_{n,m} = \frac{(n + 1)m(\frac{1}{2})^m}{m!(n + \frac{3}{2})^m}, \quad m = 0, \ldots, M \]

by recursion requires \( \mathcal{O}(M) \) operations. Since

\[
|\hat{f}_n^{[M]} - \hat{f}_n| \leq c_n (2\rho)^n \max_{k \geq 0} |\hat{v}_k| \sum_{m=M+1}^{\infty} |\chi_{n,m}|\rho^{2m},
\]

\[
|\hat{v}_n| \leq \frac{\|f\|_{L_\infty(B_\rho)}}{2\pi} \int_{-\pi}^{\pi} |1 - \rho^2 e^{2i\theta}| \, d\theta := d,
\]

we have

\[
|\hat{f}_n^{[M]} - \hat{f}_n| \leq c_n d (2\rho)^n \sum_{m=M+1}^{\infty} \rho^{2m} = \frac{c_n 2^m d}{1 - \rho^2} \rho^n + 2M + 2
\]

and this can be made arbitrarily small uniformly in \( n \) choosing sufficiently large \( M \). Hence convergence.
The strange case of $\rho \to 1$

The ellipses $B_{\rho}$ and projected roots of unity for $\rho = 0.5, 0.75, 0.95$.

As $\rho \uparrow 1$, we have

$$\hat{f}_n \to c_n \sum_{m=0}^{\infty} \chi_{n,m}[\hat{\tau}_{n+2m} - \hat{\tau}_{n+2m+2}],$$

where $\hat{\tau}_n = \pi^{-1} \int_0^\pi f(\cos \theta) \cos n\theta \, d\theta$ is the $n$th Chebyshev coefficient of $f$. This can be computed fast with Fast Cosine Transform.
2.3 Connection coefficients

The method admits an alternative explanation. Let \( \{p_n\}_{n=0}^{\infty} \) and \( \{q_n\}_{n=0}^{\infty} \) be two sets of orthogonal polynomials and \( f \) a ‘nice’ function, then

\[
 f(x) = \sum_{n=0}^{\infty} \bar{f}_n p_n(x) = \sum_{n=0}^{\infty} \hat{f}_n q_n(x),
\]

\[
 \bar{f}_n = \frac{\langle f, p_n \rangle_p}{\langle p, p \rangle_p}, \quad \hat{f}_n = \frac{\langle f, q_n \rangle_q}{\langle q, q \rangle_q}, \quad n \in \mathbb{Z}_+.
\]

The connection coefficients \( \alpha_{m,n}, 0 \leq n \leq m \) are numbers s.t.

\[
 p_m(x) = \sum_{n=0}^{m} \alpha_{m,n} q_n(x), \quad n \in \mathbb{Z}_+,
\]

and they always exist since an OPS form a basis of \( \mathbb{P} \).
Letting $p_n = T_n$ and $q_n = P_n$, it is trivial that

$$\hat{f}_n = \sum_{m=n}^{\infty} \alpha_{n,m} \tilde{f}_m, \quad n \in \mathbb{Z}_+$$

and, truncating the sum, this affords us a way to extend Fast Chebyshev Transform to Legendre (and to other OP expansions). The connection coefficients can be derived directly from

$$x^n = \frac{n!}{2^n} \sum_{k=0}^{[n/2]} \frac{(2n - 4k + 1)P_{n-2k}(x)}{k!\left(\frac{3}{2}\right)_{n-k}},$$

$$T_m(x) = \delta_{m,0} + \frac{m}{2} \sum_{k=1}^{[m/2]} \frac{(-1)^k(m - k - 1)!2^{m-2k}}{k!(m - 2k)!} x^{m-2k},$$

except that the outcome is precisely the previous method for $\rho \uparrow 1$.

Does it mean that we don’t need anymore the convoluted hypergeometric expansion? Not really! Provided $f$ is smooth in the complex plane, convergence occurs much faster for $\rho < 1$ – actually, the smaller $\rho$, the faster the convergence.
The approach of hypergeometric identities has been extended to ultraspherical polynomials.

The approach of connection coefficients has been generalised to all Jacobi polynomials.

Can we extend it to all OP expansions? E.g. to Laguerre or Hermite expansions?

No! The entire idea is that the connection coefficients decay very rapidly for any fixed $n$ and growing $m$ so that for every $\varepsilon > 0$ there exists (hopefully small) $M$ s.t.

$$\sum_{m=n+M+1}^{\infty} |\alpha_{n,m}| < \varepsilon, \quad n \in \mathbb{Z}_+,$$

because this ensures convergence.
To illustrate this point, consider the growth (or otherwise) of connection coefficients from Legendre to Hermite polynomials,

\[ \alpha_{2n,2m} = \sum_{k=0}^{n-m} (-1)^k \binom{2n}{k} \frac{2n + 2m + 1}{(n - k - m)! \left(\frac{3}{2}\right)_{n-k+m}}, \quad 0 \leq m \leq n \]

(odd coefficients behave similarly) and denote by \( \kappa_\varepsilon \in \mathbb{N} \) the least index s.t. \( |\alpha_{2n,2m}| < \varepsilon \) for all \( m > \kappa_\varepsilon \):

The function is linear, hence the cost of an algorithm to convert Legendre to Hermite expansion is \textit{quadratic} in \( n \) – and this is unacceptable!
How does this compare with our Fast Legendre Transform?

Once we form adaptive sums

$$\hat{f}_n \approx (2\rho)^n c_n \sum_{m=0}^{M_n} \chi_{n,m} \rho^{2m} \hat{v}_{n+2m},$$

where

$$M_n = \max\{a + bn, 0\}$$

for suitable values of $a, b$, the summation cost just $O(N)$ operations and the main expense is the cost of FFT.
2.4 Fast Legendre Transform

via perturbed Chebyshev grid

We consider two problems, one the (scaled) inverse of the other:

**Legendre to function (LtF):** Given \( c_0, c_1, \ldots, c_{N-1} \), compute

\[
f_k = \sum_{n=0}^{N-1} c_n P_n(\xi_k^{(N)}), \quad k = -, \ldots, N - 1,
\]

where \( \xi_{N-1}^{(N)} < \xi_{N-2}^{(N)} < \cdots < \xi_0^{(N)} \) are the zeros of \( P_N \);

**Function to Legendre (FtL):** Given \( b_0, \ldots, b_{N-1}, f_0, \ldots, f_{N-1} \), compute

\[
c_n = \left( n + \frac{1}{2} \right) \sum_{k=0}^{N-1} b_k f_k P_n(\xi_k^{(N)}), \quad n = 0, \ldots, N - 1.
\]

Choosing \( b_0, \ldots, b_{N-1} \) as the weights of Gauss–Legendre quadrature (and bearing in mind that \( \xi_0^{(N)}, \ldots, \xi_{N-1}^{(N)} \) are its nodes), FtL is a means to compute Legendre coefficients.
Let $\xi_k^{(N)} = \cos \theta_k$, $k = 0, \ldots, N - 1$. We first describe the Chebyshev to function (CtF) summation

$$f_k = \sum_{n=0}^{N-1} c_n T_n(\xi_k^{(N)}) = \sum_{n=0}^{N-1} c_n \cos(n\theta_k), \quad k = 0, \ldots, N - 1,$$

in $\mathcal{O}(N \log N)$ operations. Set $\theta_k = \theta^*_k + \delta_k$, where $\theta^*_k = \left( k + \frac{1}{2} \right) \pi / N$ is the $k$th Chebyshev point. Then

$$\cos(n\theta_k) = \sum_{\ell=0}^{\infty} (-1)^\ell \frac{(n\delta_k)^{2\ell}}{(2\ell)!} \cos(n\theta^*_k) - \sum_{\ell=0}^{\infty} (-1)^\ell \frac{(n\delta_k)^{2\ell+1}}{(2\ell+1)!} \sin(n\theta^*_k).$$

Truncating the series,

$$f_k \approx \sum_{\ell=0}^{M} (-1)^\ell \frac{\delta_k^{2\ell}}{(2\ell)!} \sum_{n=0}^{N-1} c_n n^{2\ell} \cos \frac{n(1 + \frac{1}{2})\pi}{N}$$

$$- \sum_{\ell=0}^{M} (-1)^\ell \frac{\delta_k^{2\ell+1}}{(2\ell+1)!} \sum_{n=0}^{N-1} c_n n^{2\ell+1} \sin \frac{n(1 + \frac{1}{2})\pi}{N}.$$
This ‘costs’ $M + 1$ Fast Cosine Transforms and $M + 1$ Fast Sine Transforms, at the combined price tag of $O(MN \log N)$ operations and makes sense only once $M$ is small – i.e. if the $\delta_k$'s are small.

And this is the case! Indeed, $|\delta_k| \leq \frac{\pi}{2(N+1)}$ (and much stronger bounds).
**LtC**

We need to ‘translate’ from Chebyshev-to-function into Legendre-to-function by a Legendre-to-Chebyshev transform: given coefficients $c_0, \ldots, c_{N-1}$, find $\hat{c}_0, \ldots, \hat{c}_{N-1}$ s.t.

$$\sum_{n=0}^{N-1} c_n P_n(\cos \theta_k) = \sum_{n=0}^{N-1} \hat{c}_n T_n(\cos \theta_k), \quad k = 0, \ldots, N - 1.$$  

Let $\theta = [\theta_0, \theta_1, \ldots, \theta_{N-1}]^\top$, then there exist $N \times N$ matrices $C_N$ and $P_N$ so that for a polynomial $p \in \mathbb{P}_{N-1}$

$$p_N(\theta) = T_N(\theta)\hat{c} = P_N(\theta)c \quad \Rightarrow \quad \hat{c} = T_{N}^{-1}(\theta)P_N(\theta)c.$$  

Computing $T_{N}^{-1}(\theta)d$ is precisely CtF and we can do it in $O(N \log N)$ operations – the challenge is to sum $d = P_N(\theta)c$. 
Asymptotically,

\[ P_n(\cos \theta) \approx c_n \sum_{m=0}^{M-1} \alpha_{n,m} \frac{\cos((n\theta - (m + \frac{1}{2})(\frac{\pi}{2} - \theta)))}{(2 \sin \theta)^{m+\frac{1}{2}}} , \]

where

\[ c_n = \frac{3}{\pi^{1/2}} \frac{n!}{(\frac{3}{2})_n} , \quad \alpha_{n,m} = \frac{[(\frac{1}{2})_m]^2}{m!(n + \frac{3}{2})_n} , \quad m = 0, \ldots, M - 1. \]

Therefore

\[ P_n(\cos \theta) \approx c_n \sum_{m=0}^{M-1} \left[ u_m(\theta) U_{n-1}(\cos \theta) + v_m(\cos \theta) T_n(\cos \theta) \right] , \]

where

\[ u_m(\theta) = \frac{\sin \theta \sin((m + \frac{1}{2})(\frac{\pi}{2} - \theta)))}{(2 \sin \theta)^{m+\frac{1}{2}}} , \quad v_m(\theta) = \frac{\cos((m + \frac{1}{2})(\frac{\pi}{2} - \theta)))}{2(\sin \theta)^{m+\frac{1}{2}}} . \]
Let $\theta^C$ be the vector of Chebyshev points. We partition

$$P_N(\theta^C) = P^\text{REM}_N(\theta^C) + \sum_{k=1}^{K} P_N^{[k]}(\theta^C),$$

where, setting the tolerance $\varepsilon > 0$,

$$n_M = \left[ \frac{1}{2} \left\{ \left[ \frac{1}{2} \right]_M \right\}^2 \right]^\frac{1}{M+\frac{1}{2}} \left\{ \frac{1}{\pi^{3/2} M! \varepsilon} \right\},$$

$$\left( P_N^{[k]}(\theta^C) \right)_{i,j} = \begin{cases} (P_N(\theta^C))_{i,j}, & i_k \leq i \leq N - i_k, \ \alpha^k N \leq j \leq \alpha^{k-1} N, \\ 0, & \text{otherwise} \end{cases},$$

$$i_k = \left\lfloor \frac{N + 1}{\pi} \arcsin \frac{n_M}{\alpha^k N} \right\rfloor, \quad \alpha = \min \left\{ \frac{1}{2}, \frac{1}{\log \frac{N}{n_M}} \right\},$$

$$\left( P^\text{REM}_N(\theta^C) \right)_{i,j} = P_N(\theta^C) - \sum_{k=1}^{K} P_N^{[k]}(\theta^C).$$
The product $P_N^{\text{REM}} c$ can be done in $O(KN/\alpha)$ operations, since it has 
$\approx \frac{2}{\pi} KN (\alpha^{-1} - 1)$ nonzero terms and $P_N^{[k]}$ can be computed in $O(KN/\alpha)$ 
operations. Balancing the two gives $\alpha$. Moreover, requiring $\alpha^{K+1} N < n_M$, 
we have $K = O(\log N / \log \log N)$. The total cost is $O\left(N (\log N)^2 / \log \log N\right)$.  

An obvious idea is to use Clenshaw–Curtis quadrature. Thus,

\[ \hat{f}_n = \left( n + \frac{1}{2} \right) \int_{-1}^{1} f(x) P_n(x) \, dx \]

\[ = \left( n + \frac{1}{2} \right) \int_{0}^{\pi} f(\cos \theta) P_n(\cos \theta) \sin \theta \, d\theta. \]

The function \( F_n(\theta) = f(\cos \theta) P_n(\cos \theta) \) is even and periodic and

\[ \hat{f}_n = \tilde{f}_{n,0} + 2 \sum_{\ell=1}^{\infty} \frac{\tilde{f}_{n,\ell}}{1 - 4\ell^2}, \]

where the \( \tilde{f}_{n,\ell} \)s are the Chebyshev coefficients of \( F_n \),

\[ \tilde{f}_{n,\ell} = \frac{2}{\pi} \int_{0}^{\pi} F_n(\theta) \cos \ell \theta \, d\theta, \quad \ell, n \in \mathbb{Z}_+. \]

This can be approximated by a Discrete Cosine Transform.
\[ f_{2\ell} \approx n + \frac{1}{2} \left\{ \frac{(-1)^\ell f(-1) + f(1)}{2} + (-1)^\ell f(0) P_n(0) \right\} \]

\[ + \sum_{m=1}^{N-1} \left[ f(\cos \frac{\pi m}{2N}) + (-1)^n f(-\cos \frac{\pi m}{2N}) \right] P_n(\cos \frac{\pi m}{2N}) \cos \frac{\pi m \ell}{N} \].

In other words, just a single FCT (at the cost of \( \mathcal{O}(N \log N) \)), plus some additional \( \mathcal{O}(N) \) housekeeping.

Similar approach can be extended to other ‘nice’ weights living in \((-1, 1)\), e.g. Jacobi polynomials.
Let \( G(x, y) = \log(x - y) \), where \( v = v_1 + iv_2 \). Since

\[
\log |x - y| = \text{Re} \log(x - y),
\]

it is (up to a factor of \(-1/(2\pi)\), which we suppress) the complex representation of the fundamental solution of \( \Delta^2 u = 0 \) in \( \mathbb{R}^2 \). We wish to compute

\[
u_m = \sum_{j=1}^{N} G(x_m, x_j)q_j, \quad m = 1, \ldots, N,
\]

where the \( u_m \)'s are the potentials and \( q_j \)'s the sources, while \( M, N \gg 1 \).

Naive computation takes \( O(MN) \) operations: can we do better?
Well-separated sources and targets

Let $q_j = q(y_j)$, $u_m = u(x_m)$ and suppose that the $x_m$s and $y_j$s are separated.
We have

$$\log(x - y) = \log(x - c)1 - \sum_{p=1}^{\infty} \frac{1}{p} \left(\frac{y - c}{x - c}\right)^p,$$

with rapid convergence because $|y - c| \ll |x - c|$. Therefore

$$u_m = u(x_m) = \sum_{j=1}^{N} \log(x_m - y_j)q_j$$

$$\approx \log(x_m - c)\tilde{q}_0 - \sum_{p=1}^{P} \frac{1}{(x_m - c)^p}\tilde{q}_p,$$

where

$$\tilde{q}_0 = \sum_{j=1}^{N} q_j, \quad \tilde{q}_p = -\frac{1}{p} \sum_{j=1}^{N} (y_j - c)^p q_j, \quad p = 1, \ldots, P.$$

Computing the $\tilde{q}_p$ takes $O(NP)$ operations and forming the $u_m$s $O(MP)$: altogether the operations’ count is $O((M + N)P)$ and note that $P \ll M, N$. 36
The Fast Multipole Method

Divide the computational domain into boxes,

Compute the interaction between particles in a box and particles in well-separated boxes (far-field potentials) by the above procedure, the rest (near-field potentials) by direct calculation.
1. The procedure itself is considerably more complicated and awash with detail – and its programming is a nightmare.

2. The overall cost is $O(N \log N)$ (for $M \approx N$) but further refinement reduces it to $O(N)$.

3. The procedure can be nested and is amenable to parallelisation.

4. FMM can be extended to 3D (but not beyond!) but matters become even more complicated.

5. The underlying issue is a matrix-times-vector problem and FMM can be rephrased as a matrix tessalation: this point of view will be useful in the next section.

**Main applications of FMM:** Solving the Poisson equation in nice geometries in 2D and 3D; solving the Helmholtz equation $\nabla^2 u + \lambda u = 0$; solving integral equations.
2.5 FMM applied to fast transforms

We revisit the problem of computing the Legendre transforms

\[ f(x) \approx \sum_{n=0}^{N-1} c_n P_n(x) = \sum_{j=0}^{N-1} d_j T_j(x). \]

Let \( d = M^N c \), hence \( c = L^N d \), \( L = M^{-1} \). It is easy to verify that \( M \) and \( L \) are upper triangular. Their coefficients can be calculated explicitly.

\[ \Lambda(z) = \frac{\Gamma(z + \frac{1}{2})}{\Gamma(z + 1)}, \quad \mathcal{M}(x, y) = \frac{2}{\pi} \Lambda\left(\frac{y-x}{2}\right) \Lambda\left(\frac{y+x}{2}\right), \]

\[ \mathcal{L}(x, y) = -\frac{(x + \frac{1}{2})y}{(y + x + 1)(y - x)} \Lambda\left(\frac{y-x-2}{2}\right) \Lambda\left(\frac{y+x-1}{2}\right), \]

\[
(M^N)_{i,j} = \begin{cases} 
\mathcal{M}(i, j), & 0 < i \leq j < N, i + j \text{ even,} \\
0, & \text{otherwise,}
\end{cases}
\]

\[
(L^N)_{i,j} = \begin{cases} 
\mathcal{L}(i, j), & 0 \leq i \leq j < N, i + j \text{ even,} \\
0, & \text{otherwise.}
\end{cases}
\]
Recall the well separated points from the FMM theory: specifically, given the square

\[ S = \{(x, y) : x_0 \leq x \leq x_0 + a, \ y_0 \leq y \leq y_0 + a\}, \]

we say that \( S \) is separated from the diagonal if \( y_0 - x_0 \geq 2a \).

Likewise, if \( A \) is an \( N \times N \) upper triangular matrix, set

\[ T_{i,j} = A_{p+i, q+j}, \quad i, j = 1, \ldots, m, \]

where \( p, q \geq 0 \) are given. We say that the submatrix \( T \) is separated from the diagonal (SFD) if \( q - p \geq 2m \).

Let \( T \) be an \( m \times m \) SFD submatrix of \( M^N \) (similar analysis applies to \( L^N \)) and consider the product \( w = Tv \).
Let \( t_i \) be the \( i \)th Chebyshev point and
\[
\ell_i(x) = \prod_{\substack{j=0 \atop j \neq i}}^{N-1} \frac{x - t_j}{t_i - t_j},
\]
the \( i \)th Lagrange’s interpolation polynomial. Then the error of the approximation at Chebyshev points,
\[
M(i_0 + i, j_0 + j) \approx \sum_{r=0}^{k-1} M(i_0 + i, j_0 + trm) \ell_r \left( \frac{j}{m} \right),
\]
decays like \( O(3^{-k}) \). Therefore
\[
w_i = \sum_{j=0}^{m-1} M(i_0 + i, j_0 + j) v_j \approx \sum_{r=0}^{m-1} M(i_0 + i, j_0 + trm) b_r,
\]
where
\[
b_r = \sum_{j=0}^{m-1} \ell_r \left( \frac{j}{m} \right) v_j, \quad r = 0, \ldots, k - 1.
\]
All this can be calculated in $O(km)$ operations. The trick is now to tile the upper triangle of $M^N$ with SFD submatrices, calculating the remaining ‘junk’ directly.
• The cost of the entire operation scales like $\mathcal{O}(N \log N)$.
• Coding is very complicated but, fortunately, there exist off-the-shelf programs.
• It is possible to generalise this approach to a larger set of OPs, e.g. Jacobi polynomials.
References

Spectral methods:

Fast Legendre Transform via hypergeometric identities:
Connection coefficients:

Fast Legendre Transform via perturbed Chebyshev grid:

Fast Multipole Method:


FMM applied to fast transforms: