

Diagonalising Infinite-Dimensional Operators

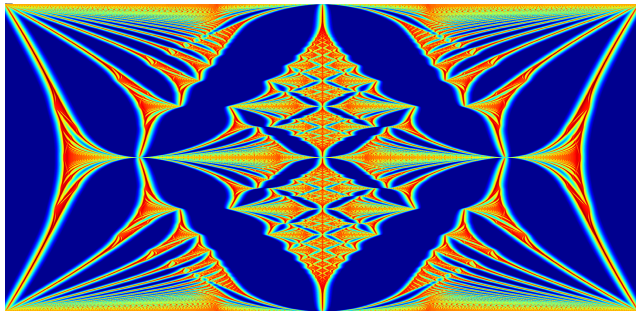
Computing spectral measures of self-adjoint operators

Matthew Colbrook
University of Cambridge

AustMS 2020

- Papers:** (i) C., Horning, Townsend, "Computing spectral measures of self-adjoint operators", under revision at *SIAM Review*.
(ii) C., "Computing spectral measures and spectral types"

Software: SpecSolve available at <https://github.com/SpecSolve>



Spectral Measures

Finite-dimensional: $A \in \mathbb{C}^{n \times n}$ self-adjoint, o.n. basis of e-vectors $\{v_j\}_{j=1}^n$

$$v = \left(\sum_{k=1}^n v_k v_k^* \right) v, \quad v \in \mathbb{C}^n \quad Av = \left(\sum_{k=1}^n \lambda_k v_k v_k^* \right) v, \quad v \in \mathbb{C}^n.$$

Infinite-dimensional: Self-adjoint operator $\mathcal{L} : \mathcal{D}(\mathcal{L}) \rightarrow \mathcal{H}$ with spectrum

$$\Lambda(\mathcal{L}) = \{z \in \mathbb{C} : \mathcal{L} - z \text{ not bounded invertible}\}.$$

Bad news: Typically, no longer an o.n. basis of e-vectors.

Spectral Theorem: Projection-valued spectral measure \mathcal{E} (assigns an orthogonal projector to each Borel-measurable set) with

$$f = \left(\int_{\mathbb{R}} d\mathcal{E}(y) \right) f, \quad f \in \mathcal{H} \quad \mathcal{L}f = \left(\int_{\mathbb{R}} y d\mathcal{E}(y) \right) f, \quad f \in \mathcal{D}(\mathcal{L}).$$

Intuition: Diagonalises an infinite-dimensional operator.

GOAL: Compute (scalar versions of) \mathcal{E} .

Motivation

Scalar-valued measures (action of projections):

$$\mu_f(\Omega) = \langle \mathcal{E}(\Omega)f, f \rangle$$

Lebesgue decomposition theorem:

$$d\mu_f(y) = \underbrace{\sum_{\lambda \in \Lambda^p} \langle \mathcal{P}_\lambda f, f \rangle \delta(y - \lambda) dy}_{\text{discrete part}} + \underbrace{\rho_f(y) dy + d\mu_f^{(\text{sc})}(y)}_{\text{continuous part}}.$$

Crucial in: quantum mechanics, scattering in particle physics, correlation in stochastic processes/signal-processing, fluid stability, resonances, density-of-states in materials science, orthogonal polynomials, random matrix theory, evolution PDEs,...

Example: in quantum mechanics, μ_f describes the likelihood of different outcomes when the observable \mathcal{L} is measured. Can also solve SE

$$i \frac{df}{dt} = \mathcal{L}f, \quad f(0) = f_0, \quad \text{via} \quad f(t) = \left(\int_{\mathbb{R}} \exp(-ity) d\mathcal{E}(y) \right) f_0.$$

A Hard Problem!

"Most operators that arise in practice are not presented in a representation in which they are diagonalized... this raises the question of how to implement the methods of finite dimensional numerical linear algebra to compute the spectra of infinite dimensional operators. Unfortunately, there is a dearth of literature on this basic problem and, so far as we have been able to tell, there are no proven techniques." W. Arveson, Berkeley (1994)

Some methods do exist, but treat cases with a lot of structure (e.g. compact perturbations of tridiagonal Toeplitz, some classes of singular Sturm–Liouville operators, etc.)

In contrast, want a **general** method to resolve spectral measures of \mathcal{L} (e.g. PDEs, integral operators, infinite matrices,...) and not an underlying discretisation or truncation.

finite-dimensional NLA \Rightarrow infinite-dimensional NLA

Ideas from OPs: Computational Favard Theorem

$a_j, b_j \in \mathbb{R}$ and $a_j > 0$:

$$J = \begin{pmatrix} b_1 & a_1 & & \\ a_1 & b_2 & a_2 & \\ & a_2 & b_3 & \ddots \\ & & \ddots & \ddots \end{pmatrix}$$

OPs orthogonal w.r.t. μ_J . For $z = x + i\epsilon$, define

$$G(z) := \langle (J - z)^{-1} \mathbf{e}_1, \mathbf{e}_1 \rangle = \int_{\mathbb{R}} \frac{d\mu_J(x)}{x - z} = \frac{1}{-z + b_1 - \frac{a_1^2}{-z + b_2 - \dots}}.$$

Then have weak convergence:

$$\lim_{\epsilon \downarrow 0} \frac{\operatorname{Im}(G(z))}{2\pi} = \mu_J.$$

Ideas from Physics: Smoothed Measures

Idea: For $z = x + i\epsilon$, use

$$\mu_f^\epsilon(x) = \left\langle \frac{(\mathcal{L} - z)^{-1} - (\mathcal{L} - \bar{z})^{-1}}{2\pi i} f, f \right\rangle = \frac{1}{\pi} \int_{\Lambda(\mathcal{L})} \frac{\epsilon}{(x - \lambda)^2 + \epsilon^2} d\mu_f(\lambda).$$

Convolution with Poisson kernel: smoothed measure.

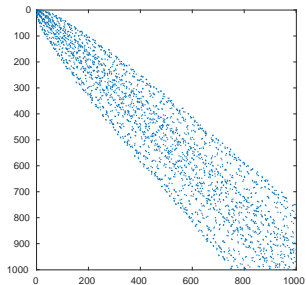
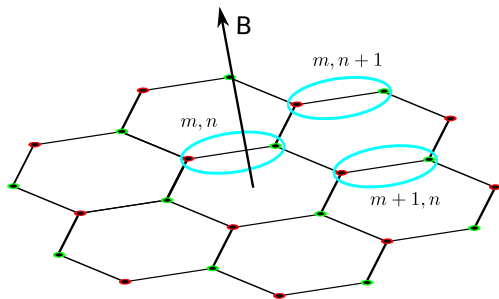
Converges weakly to measure as $\epsilon \downarrow 0$:

$$\int_{\mathbb{R}} \phi(y) \mu_f^\epsilon(y) dy \rightarrow \int_{\mathbb{R}} \phi(y) d\mu_f(y), \quad \text{as } \epsilon \downarrow 0,$$

for any bounded, continuous function ϕ .

Approximate μ_f^ϵ via $\mu_{f,N}^\epsilon$ ($N = \text{truncation parameter}$).

Numerical Balancing Act: Magnetic Graphene



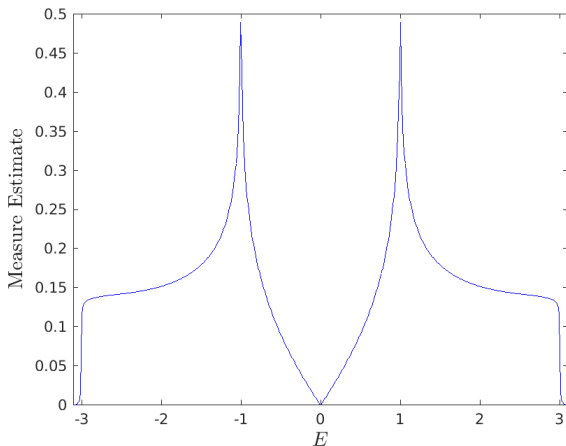
Numerical Balancing Act: Magnetic Graphene

Numerical Balancing Act: Magnetic Graphene

Theorem (C. Preprint (2019))

If we know rate of off-diagonal decay of infinite matrix, can compute measure in one limit. Extends to other operators such as PDEs.

This is through a rectangular least squares type problem that computes $(\mathcal{L} - z)^{-1}f$ with (asymptotic) error control. $N(\epsilon)$ chosen **adaptively**.

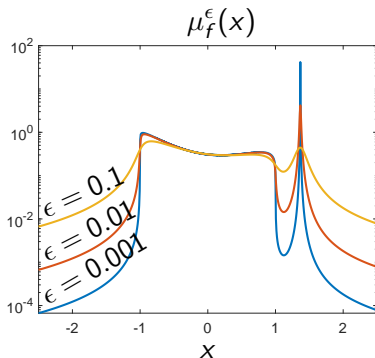


Example: Integral Operator

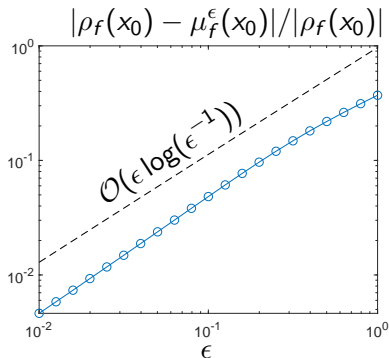
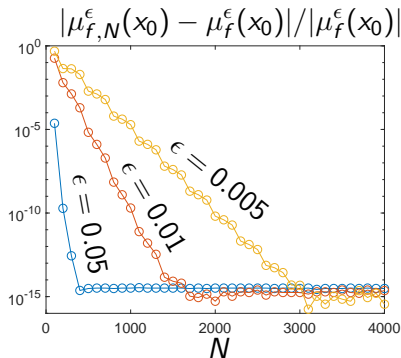
$$\mathcal{L}u(x) = xu(x) + \int_{-1}^1 e^{-(x^2+y^2)} u(y) dy, \quad x \in [-1, 1].$$

Discretise using adaptive Chebyshev collocation method.

Look at μ_f with $f(x) = \sqrt{3/2}x$.



Example: Integral Operator



$|\rho_f(x_0) - \mu_f^\epsilon(x_0)| = \mathcal{O}(\epsilon \log(\epsilon^{-1}))$ and need $N \approx 20/\epsilon$.

\Rightarrow Infeasible to get more than five or six digits!

Q: Can we do better?

Accelerating Convergence

Let $m \in \mathbb{N}$, $K \in L^1(\mathbb{R})$. We say K is an m th order kernel if:

- (i) Normalized: $\int_{\mathbb{R}} K(x) dx = 1$,
- (ii) Zero moments: $K(x)x^j$ integrable, $\int_{\mathbb{R}} K(x)x^j dx = 0$ for $0 < j < m$,
- (iii) Decay at $\pm\infty$: There is a constant C_K , independent of x , such that

$$|K(x)| \leq C_K(1 + |x|)^{-(m+1)}, \quad x \in \mathbb{R}.$$

Theorem (C., Horning, Townsend. Preprint (2020))

If K is m th order, $K_{\epsilon}(x) = \epsilon^{-1}K(x\epsilon^{-1})$ and μ_f locally absolutely continuous near x_0 with density ρ_f then

- Pointwise: If ρ_f locally $\mathcal{C}^{n,\alpha}$ near x_0 then

$$|[K_{\epsilon} * \mu_f](x_0) - \rho_f(x_0)| = \mathcal{O}(\epsilon^{n+\alpha}) + \mathcal{O}(\epsilon^m \log(\epsilon^{-1}))$$

- L^p : If ρ_f locally $\mathcal{W}^{n,p}$ near x_0 ($1 \leq p < \infty$) then

$$\|[K_{\epsilon} * \mu_f] - \rho_f\|_{L^p_{loc}} = \mathcal{O}(\epsilon^n) + \mathcal{O}(\epsilon^m \log(\epsilon^{-1}))$$

Rational Kernels

Idea: Replace Poisson kernel with rational kernel

$$K(x) = \frac{1}{2\pi i} \sum_{j=1}^m \frac{\alpha_j}{x - a_j} - \frac{1}{2\pi i} \sum_{j=1}^m \frac{\beta_j}{x - b_j}.$$

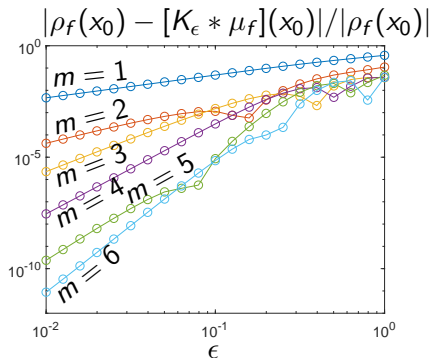
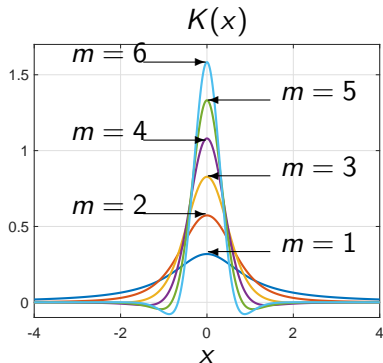
Can compute convolution with error control using resolvent

$$\begin{aligned} & [K_\epsilon * \mu_f](x) \\ &= \frac{-1}{2\pi i} \left[\sum_{j=1}^m \alpha_j \langle (\mathcal{L} - (x - \epsilon a_j))^{-1} f, f \rangle - \sum_{j=1}^m \beta_j \langle (\mathcal{L} - (x - \epsilon b_j))^{-1} f, f \rangle \right]. \end{aligned}$$

Fix a_j in UHP, b_j in LHP \Rightarrow unique $\{\alpha_j, \beta_j\}$ s.t. K an m th order kernel.

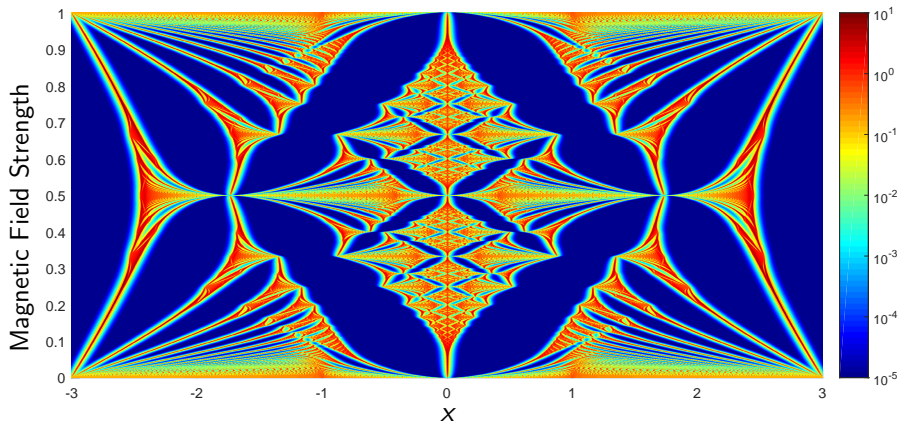
NB: At moment recommend $\{a_j = \overline{b_j}\}$ equally spaced along $\{\text{Im}(z) = 1\}$.

Integral Operator Revisited



See paper for general differential (even PDEs), integral and lattice operator examples - use sparse spectral methods for discretisation.

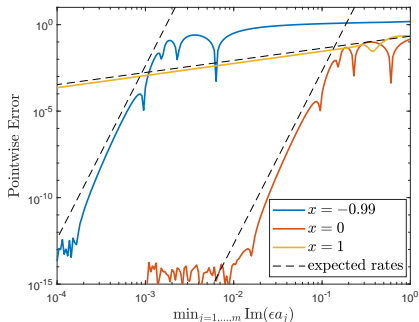
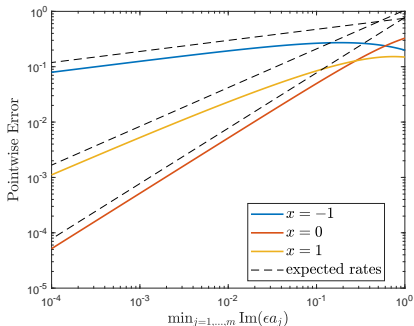
Beautiful Fractal Structure!



Spectral measure of magnetic graphene, computed to high precision (see log scale) using $m = 4$ kernel.

OP Example: Jacobi Polynomials

$$d\mu_J = \frac{(1-x)^\alpha(1+x)^\beta}{N(\alpha,\beta)} dx = f_{\alpha,\beta}(x) dx,$$



Left: Pointwise errors for $x = -1, 0, 1$ for $m = 1$ and $\alpha = 0.7$, $\beta = 0.3$. Right: Pointwise errors for $x = -0.99, 0, 1$ for $m = 10$ and $\alpha = 0.7$, $\beta = -0.3$.

Eigenvalue Hunting

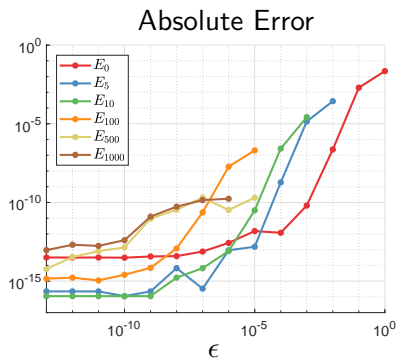
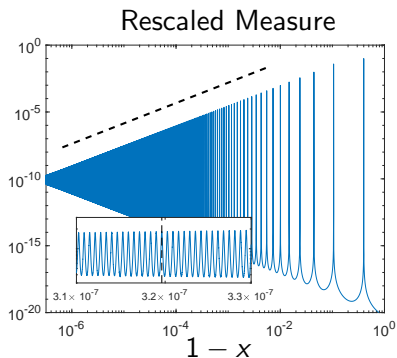
Example: Dirac operator.

- Describes the motion of a relativistic electron.
- Essential spectrum given by $\mathbb{R} \setminus (-1, 1) \Rightarrow$ spectral pollution!
- Consider radially symmetric potential, coupled system on half-line:

$$\mathcal{D}_V = \begin{pmatrix} 1 + V(r) & -\frac{d}{dr} + \frac{\kappa}{r} \\ \frac{d}{dr} + \frac{\kappa}{r} & -1 + V(r) \end{pmatrix}.$$

- Map to $[-1, 1]$ and solve shifted linear systems using sparse spectral methods.

Eigenvalue Hunting



NB: Previous state-of-the-art achieves a few digits for a few excited states.

Programme: Foundations of Infinite-Dimensional Spectral Computations

Key Question: What is possible in infinite-dimensional NLA?

How: Deal with operators directly, instead of previous 'truncate-then-solve'

⇒ Compute many spectral properties for the first time.

Framework: Classify problems in a computational hierarchy measuring their intrinsic difficulty and the optimality of algorithms.¹

⇒ Algorithms that realise the boundaries of what computers can achieve.

Have foundations for: spectra with error control, spectral type (pure point, absolutely continuous, singularly continuous), Lebesgue measure and fractal dimensions of spectra, discrete spectra, essential spectra, eigenvectors + multiplicity, spectral radii, essential numerical ranges, geometric features of spectrum (e.g. capacity), spectral gap problem, ...

¹Holds regardless of model of computation (Turing, analog,...).

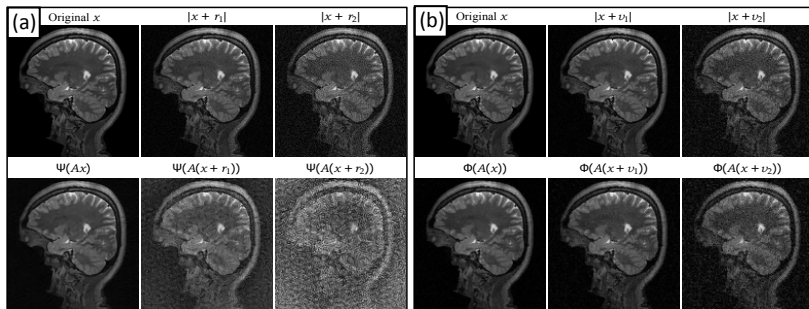
Concluding Remarks

- **DIAGONALISATION:** General framework for computing spectral measures of self-adjoint operators.
- **Convolution with RATIONAL KERNELS:**
 - Can be evaluated using resolvent. ALL you need to be able to do is solve linear systems and compute inner products.
 - High-order kernels \Rightarrow high-order convergence.
 - Generalises to normal operators for local spectral regions on curves.
- Fast, local and parallelisable \Rightarrow State-of-the-art results for PDEs, integral operators and discrete operators.
- Forms part of a PROGRAMME for foundations of infinite-dimensional spectral computations.

Code: <https://github.com/SpecSolve> (written with Andrew Horning).

Extending the Framework to Neural Networks

Antun, C., Hansen, “Can stable and accurate neural networks be computed? - On the barriers of deep learning and Smale’s 18th problem”



Left: Instability of AUTOMAP (Zhu et. al. *Nature*, 2018). Right: Stability of **F**ast **I**terative **R**estarted **N**ETworks (FIRENETs), proven to be accurate and also stable against adversarial attack.

Establishing the foundations of computation (boundaries of what can be achieved) in deep learning and AI: a fundamental problem for the 21st century.

References

- M.J. Colbrook, B. Roman, and A.C. Hansen. “How to compute spectra with error control.” *Physical Review Letters* 122.25 (2019).
- M.J. Colbrook, A.C. Hansen. “On the infinite-dimensional QR algorithm.” *Numerische Mathematik* 143.1 (2019).
- M.J. Colbrook, A.C. Hansen. “The foundations of spectral computations via the Solvability Complexity Index hierarchy: Part I.” *arXiv preprint*.
- M.J. Colbrook. “The foundations of spectral computations via the Solvability Complexity Index hierarchy: Part II.” *arXiv preprint*.
- M.J. Colbrook. “Computing spectral measures and spectral types.” *arXiv preprint*.
- M.J. Colbrook, A. Horning, and A. Townsend. “Computing spectral measures of self-adjoint operators.” *arXiv preprint*.

For further papers in this program and numerical code:

<http://www.damtp.cam.ac.uk/user/mjc249/home.html>

If you have further ideas or problems for collaboration, please get in touch!