Splitting methods and primal-dual formulations for higher order imaging models

Luca Calatroni
Cambridge Centre for Analysis

Abstract

We present an *alternating direction implicit* method solving numerically the TV-H^{-1} flow which shows good performance when applied to various imaging problems. Some stability issues depending on the regularization of the total variation may appear, thus explicit terms need to be controlled. A primal-dual approach and its possible applications to such problems is presented as well.

1 Higher order PDEs in Image Processing

We consider the following fourth-order PDE:

$$\begin{aligned} &u_t = \Delta q, \quad q \in \partial\mathcal{E}(u) && \text{in } \Omega \times (0,\infty) \\ &u(t=0) = u_0 && \text{in } \Omega, \end{aligned} \tag{1}$$

where $\mathcal{E}$ is the total variation (TV) seminorm, $\Omega \subset \mathbb{R}^2$ is a regular domain and u_0 is a sufficiently regular initial condition. We endow (1) with Neumann boundary conditions. Element of the subdifferential $\partial\mathcal{E}$ can be characterized as:

$$q = \begin{cases} -\nabla \cdot \left(\frac{\nabla u}{|\nabla u|}\right) & \text{if } |\nabla u(x)| \neq 0 \\ 0 & \text{if } |\nabla u(x)| = 0, \end{cases} \tag{2}$$

which shows both the fourth-differential order and the strong nonlinearity in (1). Higher order equations as the one above have been used in image processing problems because of their ability of preserving the edges in the process of reconstruction, see [5]. Compared to second-order L^2-total variation flows, they also reduce the presence of unwanted artifacts when applied to *denoising* problems and show better connectivity over large distances when applied to the problem of *inpainting* where the purpose is filling the missing parts of damaged images using the information from the surrounding areas.

Numerical schemes solving efficiently and in a reliable way equation (1) may restrict the choice of the time-step size Δt used in the discretization up to $\mathcal{O}(\Delta x^4)$, where Δx denotes the step size of the spatial grid and the strong nonlinearity characterizing the elements q in (2) may add additional constraints. Our aim is dealing with these constraints and solving the problem numerically simplifying the problem into smaller parts via a *operator splitting* technique. A *convexity splitting* method solving (1)-(2) has been proposed in [5]: the TV-H^{-1} flow is therein splitted into the difference of two convex energies and the resulting numerical scheme balances the unstable behaviours coming from the nonlinear terms thanks to the presence of an additional linear diffusion term.

2 ADI splitting schemes

We are interested in performing a *dimensional* operator splitting scheme for the computation of the numerical solution of equation (1) when applied to the problem of inpainting. In particular, we focus on the application of *alternating direction implicit* (ADI) schemes which prescribe the decomposition of the operator governing

the PDE into the sum of components such that each component acts just along one direction of the space where the problem is posed (see [4]). Mixed-directions contributes are encoded by one single component. The arising numerical scheme deals implicitly with the one-directional components and explicitly with the mixed one. Its structure depends on the particular splitting and some corrections have been proposed in [3] to improve its accuracy, typically of order two. To do so, some sort is symmetry is needed: *additive multiplicative operator splitting* (AMOS) schemes seem to provide a good and stable way to solve our problem, overcoming some stability issues depending on the regularizing parameter of the elements in the subdifferential of the total variation which may affect the process of reconstruction.

Our investigation starts from the analysis of a fourth-order linear PDE known as the *biharmonic* equation where we can show that bigger time-step sizes up to $\Delta t = \mathcal{O}(\Delta x^2)$ can be considered (numerical results are presented in [2]). In order to move to the nonlinear case, we consider an equivalent version of (1), reducing the fourth order equation to two equations of order two and then applying a suitable ADI scheme after linearizing the equation in a suitable way. This last choice is crucial and may cause unstable behaviours in the numerical schemes considered, as shown explicitly in [2].

3 Primal-dual formulation with penalty term

Another alternative approach to the formulation of the problem is motivated by earlier work on numerical characterization of the elements in the sub differential of the total variation seminorm by primal-dual iteration, as described in [1]. The new problem we aim to solve turns out to be:

$$\begin{cases} u_t = \Delta q, \\ \min\limits_{u \in BV(\Omega)} \left(\sup\limits_{\mathbf{p} \in C_0^\infty(\Omega;\mathbb{R}^2),\ \|\mathbf{p}\| \leq 1} \int_\Omega \nabla \cdot \mathbf{p} u - \int_\Omega q u \right). \end{cases} \tag{3}$$

The second equation encodes the condition $q \in \partial \mathcal{E}(u)$, thus providing an equivalent formulation of the problem (1). In (3) the constraint on $\mathbf{p}$ can be relaxed by a penalty method that adds an additional term in the minimization problem which penalizes if $\|\mathbf{p}\| > 1$. A typical choice for F would be for instance $F(s) = \frac{1}{2} \max\{s, 0\}^2$. The minimization problem appearing in (3) turns out to be:

$$\min_{u \in BV(\Omega)} \sup_{\mathbf{p} \in C_0^\infty(\Omega;\mathbb{R}^2)} \left(\int_\Omega \nabla \cdot \mathbf{p} u - \frac{1}{\varepsilon} F(\|\mathbf{p}\| - 1) - \int_\Omega q u \right) \tag{4}$$

where the parameter $1 \gg \varepsilon > 0$ is small and measures the weight of the penalization. Finding the optimality conditions for both $\mathbf{p}$ and u in (4) we end up with an alternative formulation of (1) which can be solved via a quasi-Newton method with damping (see [2]). The same strategy can also be applied to other higher-order flows, such as the L^2-Wassserstein gradient flow of the total variation which has been used for density estimation and smoothing.

References

[1] M. Benning, *Singular regularization of inverse problems*, PhD thesis, University of Münster, (2011).

[2] L. Calatroni, B. Düring, C.-B. Schönlieb, *ADI splitting schemes for a fourth-order nonlinear partial differential equation from image processing*, preprint 2012.

[3] K.J. in 't Hout, B.D. Welfert, *Stability of ADI schemes applied to convection-diffusion equations with mixed derivative terms*, Appl. Numer. Math., 57, 19-35, (2007).

[4] W. Hundsdorfer, J.G. Verwer, Numerical solution of time-dependent advection-diffusion-reaction equations, Berlin: Springer-Verlag, (2003).

[5] C.-B. Schönlieb, A. Bertozzi, *Unconditionally stable schemes for higher order inpainting*, Commun. Math. Sci., 9 (2), 413-457, (2011).

Statistical inference on Lévy processes

Alberto Coca Cabrero
Cambridge Centre for Analysis

Lévy processes are a rich class of stochastic processes which form the fundamental building block for stochastic continuous-time models with jumps. They have proved to be highly useful models in many applications within mathematical finance, insurance, queueing and storage theory, biology, physics, etc. However, in these applications the data corresponds to discrete-time observations of a stochastic process. Therefore, finding consistent estimators (i.e. that converge in probability to the true value) of the underlying continuous-time process and functionals of it is of vital importance for them. It is also important to understand how fast these estimations approach the true values as the number of observations grows. These theoretical questions proved to be technically challenging which, together with the growing demand for them, explains why there were hardly any important results in the area before the last decade.

The estimators we shall present exploit the fact that it is possible to produce approximations of a Lévy process and functionals of it through approximating the characteristics that uniquely determine the process. Let us elaborate on this. A Lévy process $(L_t)_{t\geq 0}$ is a stochastic process starting at zero with stationary independent increments whose paths are almost surely right continuous with left limits. The characteristic function of such a process satisfies that $\varphi_t(u) := E[\exp(iuL_t)] = \exp(t\psi(u))$, where

$$\psi(u) := i\gamma u - \frac{\sigma^2}{2}u^2 + \int_{\mathbb{R}} \left(e^{iux} - 1 - iux1_{|x|<1}\right)\nu(dx), \qquad u \in \mathbb{R}. \tag{1}$$

The function ψ is fully determined by the Lévy triplet (γ, σ, ν) with drift-like part $\gamma \in \mathbb{R}$, volatility $\sigma \geq 0$ and jump measure ν, which is a non-negative σ-finite measure such that $\int_{\mathbb{R}} \left(1 \wedge x^2\right)\nu(dx) < \infty$. Crucially, this triplet comprises the characteristics that uniquely determine the Lévy process. Therefore we shall consider the problem of estimating the Lévy triplet of a Lévy process or functionals of it from time-discrete observations of the process. As the reader may already be expecting, the challenges arise when dealing with the jump measure.

It is a standard result in statistics that consistent estimation of parametric densities from random observations can be achieved by the maximum likelihood estimator. As a consequence of the central limit theorem and under some mild assumptions, it can be shown that the rate of convergence of these estimates is $1/\sqrt{n}$, where n is the number of observations. However, when estimating nonparametric densities the rate is slower than or equal to the parametric rate and it is equal only in some particular cases. Since the triplet in (1) involves a non-parametrised measure we face a nonparametric estimation problem and we then expect the approximation rates to be different depending on the assumptions.

Henceforth we shall assume we have observations L_{t_k} of a Lévy process in a time interval $[0, T]$ at equidistant time steps $t_k = k\Delta, k = 0, 1, \ldots, n$, for $\Delta = T/n$ fixed. Thus, defining $X_k := L_{t_k} - L_{t_{k-1}}$, we equivalently observe n independent realisations of a random variable X with characteristic function $\varphi(u) = \exp(\Delta\psi(u))$, where ψ is as in (1).

Under this setting and under a moment assumption on ν, a consistent estimator for certain functionals of a Lévy process is constructed in [1]. This estimator is constructed making use of an identity relating ν and φ. The rate of convergence of the estimator is derived and, essentially because of this identity, it depends on the assumptions on the decay of φ at infinity. It is also proved that these rates are optimal given the respective assumptions, which is not totally surprising since it is well-known that estimators based on the characteristic function can be asymptotically efficient. For the reader familiarised with inverse problems, the rates are very similar to those of that type of problems, which depend on the decay

of the characteristic function of the error involved in the problem. This connection is present as the estimator also involves deconvolution with a certain function.

Under the assumption that φ decays polinomially, the rate of convergence is the most attractive one, $1/\sqrt{n}$. The class of Lévy processes covered by this assumption includes several relevant examples such as compound Poisson, Gamma and self-decomposable processes. However, the results in [1] do not cover the estimation of some interesting functionals such as the cumulative distribution function of ν, which we denote it by N, and they do not include any limiting distributions. In [2] a consistent estimator $\hat{N}_n$ for N is constructed using a similar identity to the one mentioned above and a Donsker-type of theorem is proved under the previous assumption on φ. More concretely, it is proved that for any $\zeta > 0$ fixed there exists a centered Gaussian Borel random variable $\mathbb{G}^\varphi$ in $L^\infty_\zeta := L^\infty((-\infty, -\zeta] \cup [\zeta, \infty))$ such that the process $\sqrt{n}(\hat{N}_n - N)$ converges in law to $\mathbb{G}^\varphi$ in the space L^∞_ζ.

Note that the condition to be away from zero mainly originates from the fact that ν is a finite measure away from zero, but not necessarily when it includes it. Therefore an interesting complementary result is to understand the behaviour of $\hat{N}_n - N$ when it approaches the origin. I.e. to obtain a central limit theorem for $\hat{N}_n(t_n)$ with $|t_n| \to 0$ and the rate in n for which it happens. Yet, the result in [2] has other extensions and applications we shall try to present if time permits: as a consequence of the distributional nature of the result, valuable statistical inference procedures such as goodness of fit tests and confidence bands can be derived from it; because of these procedures and also by themselves, Donsker-type of theorems for more general functionals than just for the cumulative distribution function of ν are highly relevant extensions of the result; finally a question of particular interest in the area of statistics for stochastic processes is whether one can allow for high-frequency observation regimes $\Delta_n \to 0$.

References

[1] Neumann, M.H., Reiss, M., *Nonparametric estimation for Lévy processes from low-frequency observations*, Bernoulli **15** (1), 2009, 223–248.

[2] Nickl, R., Reiss, M., *A Donsker theorem for Lévy measures*, *J.Funct. Anal.* , http://dx.doi.org/10.1016/j.jfa.2012.08.012, 2012.

Laplace equation in a convex polygon: A numerical implementation of the weak formulation of the Fokas method

Kevin Crooks
Cambridge Centre for Analysis

Abstract

In this talk I will describe a numerical implementation of the Fokas method to solving the Laplace equation in a convex polygon. This follows work by Fokas in [1, 2] on the method and Ashon in [3] which formulates a weak variational approach in the case of a convex polygon. This method uses strongly the analyticity of the Fourier Transform, and a powerful theorem by Paley and Wiener. The computational method uses a Galerkin approach with a natural sinc basis for the Paley–Wiener space. We discuss computational issues to this approach and possible resolutions relating to its implementation. The role of rigorous analysis cannot be underestimated in resolving these computational difficulties.

1 Statement of the problem

We wish to solve numerically the Laplace equation in a polygon, with low regularity boundary data. We consider the following Dirichlet problem:

$$\Delta q = 0 \quad \text{in our polygon } \Omega \quad , \quad q = f_j \quad \text{on each edge } \Gamma_j \text{ of the polygon for } j = 1, \ldots, n. \tag{1}$$

(For definiteness, we think of $f_j \in H^1$, though as Ashton remarks in [3], lower regularity results also hold). q being harmonic and Stokes' theorem quickly leads to the *global relation* of Fokas given in [2]:

$$\sum_{j=1}^{n} \rho_j(k) = 0 \quad k \in \mathbb{C}, \tag{2}$$

where each ρ_j is an integral of a fixed 1-form W along the j-th edge of our polygon. These ρ_j are called *spectral functions* and store information about the boundary data f_j, from which the solution (where it exists) can be determined. In [1], Fokas shows that a solution to (2) is given by the sum of integrals along specific rays l_k in the complex plane: $q_z = \frac{1}{2\pi}\sum_{k=1}^{n}\int_{l_k} e^{\mathrm{i}\lambda z}\rho_k(\lambda)d\lambda$. However, these spectral functions ρ_j also contain information about the Neumann data, which must be recovered from the Dirichlet data via a Dirichlet to Neumann map which shall be denoted DN. We study a numerical implementation of this method which follows a variational approach to the global relation (2) given by Ashton in [3].

2 The Variational set up

The spectral data ρ_j contains the Fourier transform of both the Dirichlet and Neumann data. A theorem by Paley and Wiener provides analyticity with additional structure, and the power of the Fokas method lies in utilising this analyticity.

Motivated by these considerations and that the Dirichlet data is real, the natural space to use is $PW^{\sigma_j}_{sym}$, a subspace of the Paley–Wiener space $PW^{\sigma_j} = \mathcal{F}L^2\left[-\sigma_j, \sigma_j\right]$.

Writing all the transforms of Dirichlet and Neumann data in vector form $\mathbf{\Phi}^t$ and $\mathbf{\Phi}^n$ respectively ('t' for tangential), Ashton shows that the *global relation* (2) can be written in the form

$$T(\mathbf{\Phi}^n - \mathrm{i}\mathbf{\Phi}^t) = 0, \tag{3}$$

for some explicit T. That is, given $\mathbf{\Phi}^t$ we seek a vector $\mathbf{\Phi}^n$ such that $\mathbf{\Phi}^n - \mathrm{i}\mathbf{\Phi}^t$ is in $N(T)$, the null set of T.

Theorem 2.1 (Ashton [3])**.** *For each $\mathbf{\Phi} \in X_{sym} := PW^{\sigma_1}_{sym} \times \cdots \times PW^{\sigma_n}_{sym}$ the set $DN(\Phi) := \{\Phi' \in X_{sym} : \Phi + i\Phi' \in N(T)\}$ contains a singleton element of X_{sym}. Thus T defines a Dirichlet to Neumann map $DN : \Phi \in X_{sym} \to DN(\Phi) \in X_{sym}$.*

In fact by analyticity, it suffices to solve for zeros of T on any set with an accumulation point. Ashton shows that this leads to the weak form of the problem: Letting

$$a(\mathbf{\Phi}, \mathbf{\Psi}) := \Re \sum_{k=1}^{n} \int_{\gamma_k} (T\mathbf{\Phi})_k(\lambda)\overline{(T\mathbf{\Psi})_k(\lambda)} ds(\lambda) \quad \text{and} \quad l(\mathbf{\Psi}) := -\Im \sum_{k=1}^{n} \int_{\gamma_k} (T\mathbf{\Phi}^t)_k(\lambda)\overline{(T\mathbf{\Psi})_k(\lambda)} ds(\lambda), \quad (4)$$

for an *arbitrary* choice of curves γ_k which eventually coincide with the negative real axis (ensuring convergence), then we have a solution to (3) whenever $a(\mathbf{\Phi}, \mathbf{\Psi}) = l(\mathbf{\Psi}) \quad \forall \mathbf{\Psi} \in X_{sym}$. To achieve this, Ashton proves boundedness and coercivity, and an application of Lax–Milgram provides uniqueness for this problem.

3 The focus of the talk

As is usual once reaching this stage, we may choose a basis for our space, and Céa's Lemma controls the error. In the Paley–Wiener space, the natural basis to use is the sinc basis $e^j_J(\lambda) := \frac{\sin(\sigma_j\lambda - \pi J)}{\sigma_j\lambda - \pi J}$. Thus the variational problem given by Ashton becomes

$$\sum_{i=1}^{n} \sum_{|I| \le N} \mathbf{\Phi}^I_i a(e_i \otimes e^i_I, e_j \otimes e^j_J) = l(e_i \otimes e^i_I), \quad 1 \le i \le n, |I| \le N, \quad (5)$$

which can be inverted giving us the projection of the solution q_z onto our chosen finite dimensional subspace.

The numerical difficulty however, comes in computing these values for a finite basis: Assuming for simplicity that γ_k is the negative real axis, then an arbitrary term from a (which we denote by $A_{k,i,I,j,J}$, recalling also that a is a sum over the n paths γ_k), and l would be

$$\begin{aligned} A_{k,i,I,j,J} &= \Re \int_0^\infty e^{\eta_{i,j,k}t} \operatorname{sinc}(\pi I + \mu_{i,k}t) \operatorname{sinc}(\pi J + \overline{\mu_{j,k}}t)\, dt \\ l_{j,J} := l(e_j \otimes e^j_J) &= -\Im \sum_k \int_{\gamma_k} (T\Phi^t)_k(-t) e^{\mathrm{i}e^{\mathrm{i}\alpha_k}\overline{(m_k - m_j)t}} \operatorname{sinc}(\pi J + \overline{\mu_{j,k}}t)\, dt, \end{aligned} \quad (6)$$

where μ, α, η and m are determined explicitly from the geometry of the problem.

Equalities (6) will be the focus of the talk; we shall consider issues relating to the evaluation these terms computationally (intuitively the oscillations in sinc, while numerically difficult, are necessary for convergence), we will also remark on inverting the resulting matrix to solve the problem, and briefly mention the expected results for a test case. Intuitive results of this can be seen from viewing the Dirichlet problem as a layer potential (i.e. a surface of charge). From this perspective $\nabla q(x, y)$ is the potential gradient of the resulting electrostatic problem. For additional details of this point the interested reader is directed to the Layer Potentials chapter in [4].

References

[1] A. S. Fokas. Two-dimensional linear partial differential equations in a convex polygon. *R. Soc. Lond. Proc. Ser. A Math. Phys. Eng. Sci.*, 457(2006):371–393, 2001.

[2] A. S. Fokas and A. A. Kapaev. On a transform method for the Laplace equation in a polygon. *IMA J. Appl. Math.*, 68(4):355–408, 2003.

[3] A. Ashton. The spectral Dirichlet–Neumann map for Laplace's equation in a convex polygon. *ArXiv e-prints*, 2012.

[4] Gerald B. Folland. *Introduction to partial differential equations.* Princeton University Press, Princeton, NJ, second edition, 1995.

Distributed solution of the optimal power flow problem

Eoin Devane
Cambridge Centre for Analysis

Introduction

Power flow studies investigate the flow of electrical power through networks with topologies that represent real-world power grids, incorporating known power loads and losses together with unknown amounts of power generation. The classical Optimal Power Flow (OPF) problem involves minimising the total cost of power generation over all of the network buses capable of generating power, subject to satisfying all of the given loads and a collection of inequality constraints corresponding to physical limits on the system's operation. Power networks today are generally partitioned in structure, with different operators running disjoint components of the overall network. Given this structure, it is desirable to investigate approaches to solving the OPF problem that are distributed in nature, whereby each operator may solve its own regional optimization problem locally. In particular, each operator would prefer to be able not to disclose the exact operational details of its own subsystem to the other operators, who may be its rivals and competitors. Consequently, we will consider an algorithmic approach, through the method of dual decomposition, to obtaining a solution of the OPF problem while requiring only minimal information sharing between the separate regions of the network.

Problem formulation and solution

The classical OPF problem is highly non-convex, making its numerical solution inefficient. We consider instead a relaxation of the problem, introduced in [1] and rigorously justified in [2], that allows it to be brought into a form that is amenable to solution as a semidefinite program. The analysis that follows is based on [3].

We model the power system in question as a network, with nodes denoting the system buses and edges denoting power transmission lines. The system consists of dense regions labeled $1, 2, \ldots, R$, connected to each other by a net of tie-lines that join a particular bus within one region to a particular bus within another region. We suppose that regions $\mathcal{I}$ and $\mathcal{J}$ are connected by exactly $n_{\mathcal{I}\mathcal{J}}$ tie-lines. In order to model each such connection, we adjoin to $\mathcal{I}$ a single dummy generator bus by means of a flow line connecting to the bus within region $\mathcal{I}$ from which the tie-line in question emanates. We do likewise for $\mathcal{J}$. In so doing we adjoin to $\mathcal{I}$ exactly $n_{\mathcal{I}} = \sum_{\mathcal{J}=1}^{R} n_{\mathcal{I}\mathcal{J}}$ dummy generator buses, each connecting to a pre-existing bus of $\mathcal{I}$ by exactly one flow line. The augmented subsystem $\mathcal{I}$ now consists of nodes that we label arbitrarily by $1, 2, \ldots, m_{\mathcal{I}} + n_{\mathcal{I}}$. We then define the maps $\ell_{\mathcal{I}\mathcal{J}}$ to be such that the dummy buses corresponding to the $n_{\mathcal{I}\mathcal{J}}$ tie-lines joining region $\mathcal{I}$ to region $\mathcal{J}$ are indexed by $\ell_{\mathcal{I}\mathcal{J}}(r)$ for $r \in \{1, 2, \ldots, n_{\mathcal{I}\mathcal{J}}\}$. Moreover, since $n_{\mathcal{J}\mathcal{I}} = n_{\mathcal{I}\mathcal{J}}$, we can stipulate that $\ell_{\mathcal{J}\mathcal{I}}(r) = \ell_{\mathcal{I}\mathcal{J}}(r)$. Finally, we denote by $f_k^{\mathcal{I}}$ the convex local cost function of power generation at a bus k in the set $\tilde{\mathcal{G}}^{\mathcal{I}}$ of generator buses, excluding all dummy buses, within the nodal set $\{1, 2, \ldots, m_{\mathcal{I}} + n_{\mathcal{I}}\}$. Using the notation of [2], the relaxed formulation of the OPF problem that we shall consider may thus be written as

$$
\begin{aligned}
\underset{(W^{\mathcal{I}})_{\mathcal{I}=1,2,\ldots,R}}{\text{minimize}} \quad & \sum_{\mathcal{I}=1}^{R} \sum_{k \in \tilde{G}^{\mathcal{I}}} f_k^{\mathcal{I}} \left(P_{d_k}^{\mathcal{I}} + \operatorname{Tr}\left(\mathbf{Y}_k^{\mathcal{I}} W^{\mathcal{I}} \right) \right) \\
\underset{\forall \mathcal{I}, \mathcal{J}, r}{\text{subject to}} \quad & \operatorname{Tr}\left(\mathbf{Y}_{\ell_{\mathcal{I}\mathcal{J}}(r)}^{\mathcal{I}} W^{\mathcal{I}} \right) + \operatorname{Tr}\left(\mathbf{Y}_{\ell_{\mathcal{J}\mathcal{I}}(r)}^{\mathcal{J}} W^{\mathcal{J}} \right) = 0, \\
& \operatorname{Tr}\left(\bar{\mathbf{Y}}_{\ell_{\mathcal{I}\mathcal{J}}(r)}^{\mathcal{I}} W^{\mathcal{I}} \right) + \operatorname{Tr}\left(\bar{\mathbf{Y}}_{\ell_{\mathcal{J}\mathcal{I}}(r)}^{\mathcal{J}} W^{\mathcal{J}} \right) = 0, \\
& \operatorname{Tr}\left(\mathbf{M}_{\ell_{\mathcal{I}\mathcal{J}}(r)}^{\mathcal{I}} W^{\mathcal{I}} \right) - \operatorname{Tr}\left(\mathbf{M}_{\ell_{\mathcal{J}\mathcal{I}}(r)}^{\mathcal{J}} W^{\mathcal{J}} \right) = 0, \\
& W^{\mathcal{I}} \in \mathcal{R}^{\mathcal{I}},
\end{aligned}
\tag{1}
$$

where $\mathcal{R}^{\mathcal{I}}$ is a convex set containing all of the load satisfaction constraints, the physical inequality constraints, and the positive semidefiniteness requirement on the matrices $W^{\mathcal{I}}$. Here, at bus k, $P_{d_k}^{\mathcal{I}}$ denotes the known real power load, and then $P_{d_k}^{\mathcal{I}} + \operatorname{Tr}\left(\mathbf{Y}_k^{\mathcal{I}} W^{\mathcal{I}}\right)$, $Q_{d_k}^{\mathcal{I}} + \operatorname{Tr}\left(\bar{\mathbf{Y}}_k^{\mathcal{I}} W^{\mathcal{I}}\right)$, and $\operatorname{Tr}\left(\mathbf{M}_k^{\mathcal{I}} W^{\mathcal{I}}\right)$ represent the classical real power generation, reactive power generation, and squared nodal voltage magnitude variables.

We apply the method of dual decomposition to the optimization problem (1), introducing dual variables for each of the tie-line equality constraints and then adding them on to the objective function. The resulting primal optimization problem then separates into optimization problems local to each region $\mathcal{I}$, for any given values of the dual variables. The

dual variables are then updated by a separate update rule such that these also converge to their optimal value. The interconnection between these primal and dual update schemes should then drive the variables towards a saddle point of the overall primal-dual problem. The precise interconnection scheme to be considered is formulated in Algorithm 1.

Algorithm 1 Distributed dual decomposition algorithm for the OPF problem.

while $\Delta_n >$ tolerance **do**

1. Optimize the separate regional problems by

$$W^{\mathcal{I}}_{n+1} = \underset{W^{\mathcal{I}} \in \mathcal{R}^{\mathcal{I}}}{\operatorname{argmin}} \left\{ \begin{aligned} &\sum_{k \in \tilde{G}^{\mathcal{I}}} f_k^{\mathcal{I}} \left(P^{\mathcal{I}}_{d_k} + \operatorname{Tr}\left(\mathbf{Y}^{\mathcal{I}}_k W^{\mathcal{I}} \right) \right) + \sum_{\mathcal{J}=1}^{R} \sum_{r=1}^{n_{\mathcal{I}\mathcal{J}}} \Bigg(\lambda^P_{\mathcal{I}\mathcal{J}r,n} \operatorname{Tr}\left(\mathbf{Y}^{\mathcal{I}}_{\ell_{\mathcal{I}\mathcal{J}}(r)} W^{\mathcal{I}} \right) + \lambda^Q_{\mathcal{I}\mathcal{J}r,n} \operatorname{Tr}\left(\bar{\mathbf{Y}}^{\mathcal{I}}_{\ell_{\mathcal{I}\mathcal{J}}(r)} W^{\mathcal{I}} \right) \\ &\qquad + \delta_{\mathcal{I}\mathcal{J}} \lambda^V_{\mathcal{I}\mathcal{J}r,n} \operatorname{Tr}\left(\mathbf{M}^{\mathcal{I}}_{\ell_{\mathcal{I}\mathcal{J}}(r)} W^{\mathcal{I}} \right) \Bigg), \end{aligned} \right.$$

according to the condition that if given the same data (choices of $\lambda^P_{\mathcal{I}\mathcal{J}r,n}$, $\lambda^Q_{\mathcal{I}\mathcal{J}r,n}$, and $\lambda^V_{\mathcal{I}\mathcal{J}r,n}$) on multiple occasions, the optimizer should on each occasion yield the same answer. Here $\delta_{\mathcal{I}\mathcal{J}}$ takes the value 1 if $\mathcal{I} > \mathcal{J}$ and -1 if $\mathcal{I} < \mathcal{J}$.

2. Update the dual variables by

$$\begin{aligned} \lambda^P_{\mathcal{I}\mathcal{J}r,n+1} &= \lambda^P_{\mathcal{I}\mathcal{J}r,n} + \alpha \Big(\operatorname{Tr}\left(\mathbf{Y}^{\mathcal{I}}_{\ell_{\mathcal{I}\mathcal{J}}(r)} W^{\mathcal{I}}_{n+1} \right) + \operatorname{Tr}\left(\mathbf{Y}^{\mathcal{J}}_{\ell_{\mathcal{J}\mathcal{I}(r)}} W^{\mathcal{J}}_{n+1} \right) \Big), \\ \lambda^Q_{\mathcal{I}\mathcal{J}r,n+1} &= \lambda^Q_{\mathcal{I}\mathcal{J}r,n} + \alpha \Big(\operatorname{Tr}\left(\bar{\mathbf{Y}}^{\mathcal{I}}_{\ell_{\mathcal{I}\mathcal{J}}(r)} W^{\mathcal{I}}_{n+1} \right) + \operatorname{Tr}\left(\bar{\mathbf{Y}}^{\mathcal{J}}_{\ell_{\mathcal{J}\mathcal{I}(r)}} W^{\mathcal{J}}_{n+1} \right) \Big), \\ \lambda^V_{\mathcal{I}\mathcal{J}r,n+1} &= \lambda^V_{\mathcal{I}\mathcal{J}r,n} + \alpha \Big(\operatorname{Tr}\left(\mathbf{M}^{\mathcal{I}}_{\ell_{\mathcal{I}\mathcal{J}}(r)} W^{\mathcal{I}}_{n+1} \right) - \operatorname{Tr}\left(\mathbf{M}^{\mathcal{J}}_{\ell_{\mathcal{J}\mathcal{I}(r)}} W^{\mathcal{J}}_{n+1} \right) \Big), \end{aligned}$$

for $\mathcal{I} \in \{1, 2, \ldots, R\}$, $\mathcal{J} \in \{1, 2, \ldots, \mathcal{I} - 1\}$, and $r \in \{1, 2, \ldots, n_{\mathcal{I}\mathcal{J}}\}$. Update the remaining dual variables by the symmetry rule $\lambda^P_{\mathcal{J}\mathcal{I}r} = \lambda^P_{\mathcal{I}\mathcal{J}r}$, $\lambda^Q_{\mathcal{J}\mathcal{I}r} = \lambda^Q_{\mathcal{I}\mathcal{J}r}$, and $\lambda^V_{\mathcal{J}\mathcal{I}r} = \lambda^V_{\mathcal{I}\mathcal{J}r}$. Here α is a positive constant.

3. Update the residual error by $\Delta_{n+1} = \max_{\mathcal{I},\mathcal{J},r} \Big\{ \Big\| \operatorname{Tr}\left(\mathbf{Y}^{\mathcal{I}}_{\ell_{\mathcal{I}\mathcal{J}}(r)} W^{\mathcal{I}}_{n+1} \right) + \operatorname{Tr}\left(\mathbf{Y}^{\mathcal{J}}_{\ell_{\mathcal{J}\mathcal{I}(r)}} W^{\mathcal{J}}_{n+1} \right) \Big\|, \Big\| \operatorname{Tr}\left(\bar{\mathbf{Y}}^{\mathcal{I}}_{\ell_{\mathcal{I}\mathcal{J}}(r)} W^{\mathcal{I}}_{n+1} \right) + \operatorname{Tr}\left(\bar{\mathbf{Y}}^{\mathcal{J}}_{\ell_{\mathcal{J}\mathcal{I}(r)}} W^{\mathcal{J}}_{n+1} \right) \Big\|, \Big\| \operatorname{Tr}\left(\mathbf{M}^{\mathcal{I}}_{\ell_{\mathcal{I}\mathcal{J}}(r)} W^{\mathcal{I}}_{n+1} \right) - \operatorname{Tr}\left(\mathbf{M}^{\mathcal{J}}_{\ell_{\mathcal{J}\mathcal{I}(r)}} W^{\mathcal{J}}_{n+1} \right) \Big\| \Big\}$.

end while

We consider this algorithm, together with some important variants of it, and analyze the significant properties of stability and convergence by assuming the step size to be small and considering the formulation of the algorithms in continuous time. For Algorithm 1, this continuous-time system is

$$\begin{aligned} \frac{d\lambda^P_{\mathcal{I}\mathcal{J}r}}{dt} &= \hat{\alpha} \left(\operatorname{Tr}\left(\mathbf{Y}^{\mathcal{I}}_{\ell_{\mathcal{I}\mathcal{J}}(r)} W^{\mathcal{I}}(\boldsymbol{\lambda}) \right) + \operatorname{Tr}\left(\mathbf{Y}^{\mathcal{J}}_{\ell_{\mathcal{J}\mathcal{I}(r)}} W^{\mathcal{J}}(\boldsymbol{\lambda}) \right) \right), \\ \frac{d\lambda^Q_{\mathcal{I}\mathcal{J}r}}{dt} &= \hat{\alpha} \left(\operatorname{Tr}\left(\bar{\mathbf{Y}}^{\mathcal{I}}_{\ell_{\mathcal{I}\mathcal{J}}(r)} W^{\mathcal{I}}(\boldsymbol{\lambda}) \right) + \operatorname{Tr}\left(\bar{\mathbf{Y}}^{\mathcal{J}}_{\ell_{\mathcal{J}\mathcal{I}(r)}} W^{\mathcal{J}}(\boldsymbol{\lambda}) \right) \right), \\ \frac{d\lambda^V_{\mathcal{I}\mathcal{J}r}}{dt} &= \hat{\alpha} \left(\operatorname{Tr}\left(\mathbf{M}^{\mathcal{I}}_{\ell_{\mathcal{I}\mathcal{J}}(r)} W^{\mathcal{I}}(\boldsymbol{\lambda}) \right) - \operatorname{Tr}\left(\mathbf{M}^{\mathcal{J}}_{\ell_{\mathcal{J}\mathcal{I}(r)}} W^{\mathcal{J}}(\boldsymbol{\lambda}) \right) \right), \end{aligned}$$

where

$$W^{\mathcal{I}}(\boldsymbol{\lambda}) = \underset{W^{\mathcal{I}} \in \mathcal{R}^{\mathcal{I}}}{\operatorname{minimize}} \left\{ \begin{aligned} &\sum_{k \in \tilde{G}^{\mathcal{I}}} f_k^{\mathcal{I}} \left(P^{\mathcal{I}}_{d_k} + \operatorname{Tr}\left(\mathbf{Y}^{\mathcal{I}}_k W^{\mathcal{I}} \right) \right) + \sum_{\mathcal{J}=1}^{R} \sum_{r=1}^{n_{\mathcal{I}\mathcal{J}}} \Bigg(\lambda^P_{\mathcal{I}\mathcal{J}r} \operatorname{Tr}\left(\mathbf{Y}^{\mathcal{I}}_{\ell_{\mathcal{I}\mathcal{J}}(r)} W^{\mathcal{I}} \right) + \lambda^Q_{\mathcal{I}\mathcal{J}r} \operatorname{Tr}\left(\bar{\mathbf{Y}}^{\mathcal{I}}_{\ell_{\mathcal{I}\mathcal{J}}(r)} W^{\mathcal{I}} \right) \\ &\qquad + \delta_{\mathcal{I}\mathcal{J}} \lambda^V_{\mathcal{I}\mathcal{J}r} \operatorname{Tr}\left(\mathbf{M}^{\mathcal{I}}_{\ell_{\mathcal{I}\mathcal{J}}(r)} W^{\mathcal{I}} \right) \Bigg). \end{aligned} \right.$$

References

[1] X. Bai, H.Wei, K. Fujisawa, and Y. Wang, "Semidefinite programming for optimal power flow problems," *Int. J. Elect. Power Energy Syst.*, vol. 30, no. 6–7, pp. 383–392, Jul.–Sep. 2008.

[2] J. Lavaei and S. H. Low, "Zero duality gap in optimal power flow problem," *IEEE Trans. Power Syst.*, vol. 27, no. 1, pp. 92–107, Feb. 2012.

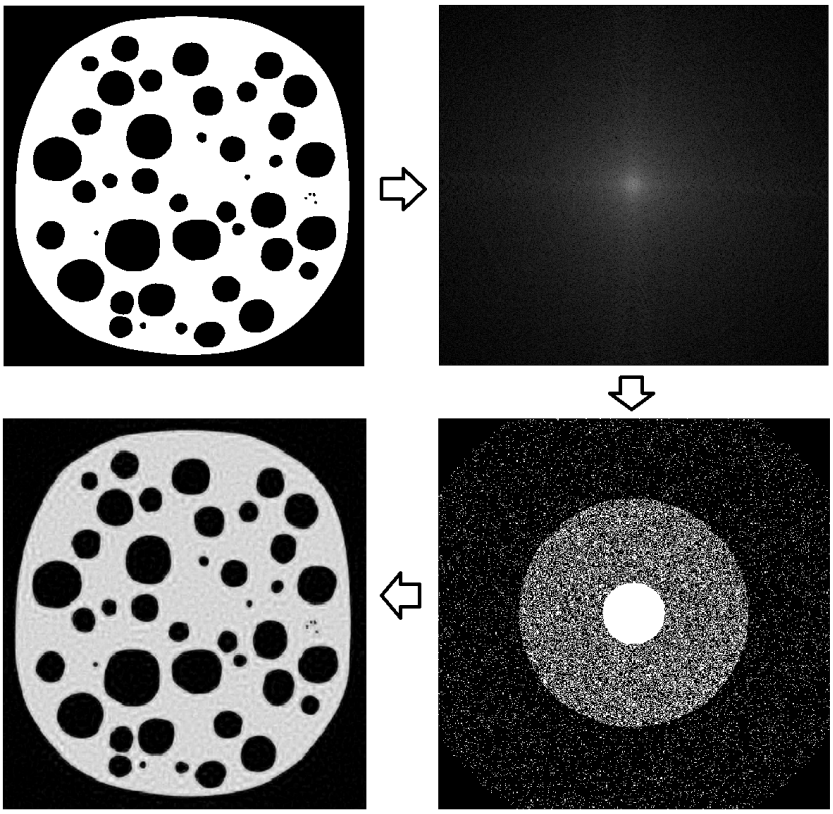
[3] E. Devane and I. C. Lestas, "Distributed solution of the OPF problem," *Preprint.*

Semirandom Sampling & Incoherence

Alex Jones
Cambridge Centre for Analysis

In NMR and MRI situations, such as trying to image a Brain or determining the structure of proteins, we are often sampling Fourier Coefficients of the Object we want to acquire. Therefore the choice of coefficients to sample is of prime importance since it effects the quality of reconstruction. The talk will present current work on finding good 1D and 2D sampling schemes based on studying the incoherence of a change of two bases (see below).

Figure 1: An example of attempted reconstruction from a sampling scheme. Going clockwise: The original phantom (rasterised), it's Fourier Coefficients, Sampling Pattern and finally Reconstruction using this pattern



Let $g \in L^2([0,1]^2)$ be the intensity of an image we would like to see. Suppose we can receive a fixed number of Fourier samples from this object, namely we can select a finite set number of points from the infinite matrix

$$S_{n,m} \quad = \quad \int_{\mathbb{R}^2} f(x,y) \exp(-2\pi i(nx+my))\, dx\, dy \qquad n,m \in \mathbb{Z}$$

Since this is too little information to determine g (there are too many potential solutions) we must use extra information about g to make our problem well posed. One possible condition is the sparsity of the

image, which means the number of non-zeros (or near zeros) of the object in question. Since our image is not typically sparse we will first transform the image in terms of a basis where it is sparse (typically a wavelet basis, but could also be another like polynomials). Therefore we think of g as a sequence of coefficients in some basis $(\psi_n)_{n\in\mathbb{N}}$ and attempt to reconstruct by solving

$$\min_{f\in\ell^2} \|f\|_1 \quad \text{subject to} \quad (Uf-b)|_S = 0$$

where b is our vector of Fourier samples, S denotes the set of subsamples and U is the change of basis matrix from our sparse basis to the Fourier basis . Note that if $(\varphi_m)_{m\in\mathbb{N}}$ denotes an ordering of the Fourier Basis $(\exp(2\pi i(nx+my)))_{m,n\in\mathbb{N}}$ then $U_{m,n} = \langle\varphi_m, \psi_n\rangle$. Typically the entries of U consist of evaluations of the Fourier Transform of some function. We will look at how the choice of S and the wavelet basis influence the quality of reconstruction.

Work by Adcock, Anders et al. (see [1]) suggests a sufficient condition for good reconstruction by breaking down both the sparse and sampling bases into regions M_k, N_k respectively. A basic form of this result states that to expect reasonable reconstruction (with error ϵ) we must subsample in N_k with percentage greater than

$$\text{Constant}(\epsilon)\cdot\mu(\perp P_{N_{k-1}}U)\cdot(\text{Sparsity}(M_1)+\cdots+\text{Sparsity}(M_k))\cdot\log(\text{Total Sparsity})$$

Where $\mu(P^{\perp}_{N_{k-1}}U)$ (called the coherence) denotes the largest absolute value in the matrix:

$$P^{\perp}_{N_{k-1}}U \;=\; \begin{pmatrix} U_{N_{k-1}+1,1} & U_{N_{k-1}+1,2} & U_{N_{k-1}+1,3} & \cdots \\ U_{N_{k-1}+2,1} & U_{N_{k-1}+2,2} & U_{N_{k-1}+2,3} & \cdots \\ U_{N_{k-1}+3,1} & U_{N_{k-1}+3,2} & U_{N_{k-1}+3,1} & \cdots \\ \vdots & \vdots & \vdots & \end{pmatrix}$$

The rest of the talk will focus on the studying the incoherence and looking at its consequences for sampling schemes in 1D and 2D.

In the 1D case with the Fourier basis as our sampling basis and wavelets as the sparse basis, the picture is very straightforward, i.e. we know the exact order of the incoherence $\mu(P^{\perp}_{N}U)$ as a function of N. The case when wavelets are replaced with polynomials is also known.

However, the 2D case is a little more complicated. There are three issues before even the problem can be posed:

1. What do we take as our wavelet basis in 2D (given a choice of 1D wavelet to build it from)?
2. What is the ordering of this wavelet basis?
3. What is the ordering of the Fourier Basis?
   (note this is not a problem in 1D since it is natural to order the frequencies $\mathbb{Z}$ by increasing size, but it is not immediately obvious what "size" should be for $\mathbb{Z}^2$)

We will look into the various results corresponding to different answers to these questions and their consequences in the search for optimal sampling patterns for a given class of images.

References

[1] B. Adcock, A. C. Hansen, C. Poon & B. Roman, *Breaking the coherence barrier: asymptotic incoherence and asymptotic sparsity in compressed sensing*, (Preprint)

Singular solutions to minimal surface equation

Kornel Maczynski
Cambridge Centre for Analysis

Minimal surfaces are the stationary points of the area functional $\mathcal{A}(S) = \mathcal{H}^n(S)$. Classical examples include plane, catenoid and helicoid. If the ambient space is a riemannian manifold and $n = 1$ then they are just the geodesics of the given manifold.

In the study of minimal surfaces it is common to use integral varifolds — a generalisation of manifolds. The Allard's regularity theorem states that, under some conditions, a minimal integral varifold is regular ($C^{1,\alpha}$) on a dense open set. Nonetheless, there are examples of minimal surfaces with singularities (e.g. a cone over $S^1 \times S^1$ in $\mathbb{R}^4$). Singularities are in fact a phenomenon that cannot be dismissed.

I would like to present findings (with possibly small contribution of my own) of the paper [1] by Caffarelli, Hardt and Simon where the authors present a procedure of obtaining examples of minimal surfaces with isolated singularities that asymptotically look like a given stationary cone.

The problem is approached using (elliptic) partial differential equations. Assuming a minimal surface is locally a graph of a ($C^{1,\alpha}$) function u then u satisfies an elliptic PDE called the minimal surface equation (MSE)

$$\nabla \cdot \left(\frac{\nabla u}{\sqrt{1 + |\nabla u|^2}} \right) = 0.$$

To follow the talk it is enough to know some basic properties of the laplacian and elliptic PDEs.

References

[1] Luis Caffarelli, Robert Hardt and Leon Simon. 'Minimal surfaces with isolated singularities'. In: *Manuscripta Math.* 48.1-3 (1984), pp. 1–18. issn: 0025-2611. doi: 10.1007/BF01168999.

A nonlinear scheme for generalized sampling

Clarice Poon
University of Cambridge, Cambridge, UK

February 18, 2013

Abstract

We consider a stable and consistent reconstruction scheme for generalized sampling and its application to wavelet reconstructions from Fourier samples. The implications of this algorithm for variable density sampling schemes in compressed sensing will also be presented.

1 Generalized sampling

A fundamental problem of signal processing is the reconstruction of signals from a discrete set of measurements. This can be formulated in a Hilbert Space $\mathcal{H}$ with inner product $\langle \cdot, \cdot \rangle$. Given any $f \in \mathcal{H}$, its measurements are of the form $\langle f, s_j \rangle$ for some orthonormal set $\{s_j\}_{j\in\mathbb{N}} \subset \mathcal{S} \leq \mathcal{H}$ and one seeks to approximate f in some orthonormal reconstruction space $\overline{\text{span}}\{w_j : j \in \mathbb{N}\} = \mathcal{W} \leq \mathcal{H}$. The goal of generalized sampling is to reconstruct in an arbitrary space $\mathcal{W}$ without placing any constraints on the type of input vectors. In practice, we seek an approximation of f in the finite dimensional space $\mathcal{W}_N = \text{span}\{w_j : 1 \leq j \leq N\}$ from some finite set of measurements $\hat{f}_M = (\langle f, s_j \rangle)_{j=1}^M$.

This problem has been extensively studied - important contributions include the consistent sampling scheme introduced by Aldroubi & Unser [6] and later extended by Eldar [3], and more recently, a reduced consistency scheme introduced by Adcock & Hansen [1]. However, the consistent sampling scheme is known to be non-convergent and ill-posed for large problem sizes in certain important cases, such as orthonormal wavelet reconstructions (other than the Haar case) from Fourier samples. Whilst the scheme of Adcock & Hansen resolves this, the resultant solution is no longer consistent with original measurements in that the reconstructed signal $R(f)$ of f from $\hat{f}_N$ could be such that $\langle R(f), s_j \rangle \neq \langle f, s_j \rangle$ for some $1 \leq j \leq N$. We will consider a non-linear scheme which presents convergent and well posed solutions to the problem of generalized sampling *and* retains the property of consistency.

2 A stable and consistent scheme

Suppose we seek to reconstruct $f \in \mathcal{W}$ such that $f = \sum_{j\in\mathbb{N}} \alpha_j w_j$. Letting $x_0 = (\alpha_j)_{j\in\mathbb{N}}$, we have that $\hat{f}_M = P_{[M]} U x_0$ where $[M] := \{1, \ldots, M\}$ and P_Ω denotes the projection onto the canonical basis indexed by the set Ω. Consider the following non-linear problem

$$\inf_{\eta\in\mathcal{H}} \|\eta\|_{l^1} \text{ subject to } P_{[M]} U \eta = P_{[M]} U x_0 \tag{1}$$

where $U = (\langle w_j, s_i \rangle)_{i,j\in\mathbb{N}}$ is the measurement matrix. Any solution to this problem will naturally be consistent with the original measurements $\hat{f}_M$. It remains to ascertain whether the solution is convergent to f as M increases. We have the following result on convergence and stability of the scheme.

Theorem 2.1. *[5] Let $x_0 \in \ell^1(\mathbb{N})$. For all $N \in \mathbb{N}$, there exists $m_0 \in \mathbb{N}$, such that for all $M \geq m_0$, if ξ solve (1), then*

$$\|\xi - x_0\|_{\ell^1} \leq C \cdot \left\| P_{[N]}^{\perp} x_0 \right\|_{\ell^1}$$

for some universal constant C. Hence, given any $f \in \mathcal{W}$ such that $f = \sum_{j=1}^{\infty} \alpha_j w_j$, if ξ solves (1) with $x_0 = (\alpha_j)_{j\in\mathbb{N}}$, then

$$\left\| f - \sum_{j=1}^{\infty} \xi_j w_j \right\| \leq C \cdot \sum_{j=N+1}^{\infty} |\alpha_j| .$$

3 Wavelet reconstructions from Fourier samples

The case where the reconstruction space $\mathcal{W}$ is generated by compactly supported orthonormal wavelets and the sampling space is the space of complex exponentials $\mathcal{S} = \overline{\text{span}} \left\{ e^{2\pi i \epsilon j \cdot} : j \in \mathbb{Z} \right\}$ for some appropriate $\epsilon > 0$ is particularly important, with applications in medical imaging and is in fact already implemented in practice [4].

The main conclusion of our analysis is: For sufficiently smooth wavelet bases, the number of reconstruction vectors which can be approximated is linearly proportional to the number of Fourier samples. Consequently, acquiring Fourier samples is up to a constant as good as acquiring the wavelet coefficients directly.

Theorem 3.1. *[5] Let the reconstruction space $\mathcal{W}$ be generated by a compactly supported Multiresolution Analysis scaling function ϕ and wavelet ψ. Suppose that ϕ and ψ are twice continuously differentiable. Then, for $N \in \mathbb{N}$, there exists some $C \in \mathbb{N}$ independent of N such that for $M = C \cdot N$, the solution ξ to (1) satisfies*

$$\|\xi - x_0\|_{\ell^1} \leq 6 \cdot \left\| P_{[N]}^{\perp} x_0 \right\|_{\ell^1} .$$

4 Consequences for compressed sensing

A mathematical framework was recently introduced in [2] to provide theoretical justifications to variable density sampling schemes often adopted when implementing compressed sensing. The theory therein contained analysis of solutions to semi-random schemes such as

$$\inf_{\eta \in \mathcal{H}} \|\eta\|_{l^1} \text{ subject to } (P_{[M]} \oplus P_\Omega) U \eta = (P_{[M]} \oplus P_\Omega) U x_0$$

where U is the measurement matrix and $\Omega \subset \{j \in \mathbb{N} : j \geq M+1\}$ is chosen uniformly at random. It is important to understand how the choice of M affects reconstruction quality, particularly for the specific case of wavelet reconstructions from Fourier samples. Understanding the error of solutions to (1) is equivalent to considering how many reconstruction coefficients can be approximated from M samples and will have implications on the choice of M in the semi-random scheme.

References

[1] B. Adcock and A. C. Hansen. A generalized sampling theorem for stable reconstructions in arbitrary bases. *J. Fourier Anal. Appl.*, 18(4):685–716, 2012.

[2] B. Adcock, A. C. Hansen, C. Poon, and B. Roman. Breaking the coherence barrier: asymptotic incoherence and asymptotic sparsity in compressed sensing. *arXiv preprint arXiv:1302.0561*, 2013.

[3] Y. C. Eldar. Sampling with arbitrary sampling and reconstruction spaces and oblique dual frame vectors. *J. Fourier Anal. Appl.*, 9(1):77–96, 2003.

[4] M. Guerquin-Kern, M. Haberlin, K. Pruessmann, and M. Unser. A fast wavelet-based reconstruction method for magnetic resonance imaging. *Medical Imaging, IEEE Transactions on*, 30(9):1649–1660, 2011.

[5] C. Poon. A stable and consistent approach to generalized sampling. *Preprint*, 2013.

[6] M. Unser and A. Aldroubi. A general sampling theory for nonideal acquisition devices. *IEEE Trans. Signal Process.*, 42(11):2915–2925, 1994.

Spectral analysis of the Fox–Li operator

Alberto Gil Couto Pimentel Ramos
Cambridge Centre for Analysis

Abstract

In this talk I present some recent work with Prof. Arieh Iserles on the asymptotic behaviour of a generic eigenvalue and eigenfunction pair of the Fox–Li operator in the asymptotic regime of large oscillations. This Fox–Li eigenvalue and eigenfunction problem is particularly relevant to the science of laser beams and resonators, as well as a very appealing theme for research given that it is (mildly) non-normal and falls outside of standard spectral theory. The talk will concern two of our different approaches and conclude with current thoughts on the problem.

1 Problem statement

The operator

$$\mathcal{F}_\omega : L^2([-1,1],\mathbb{C}) \to L^2([-1,1],\mathbb{C}) \qquad (\mathcal{F}_\omega f)(y) := \sqrt{\frac{\omega}{\pi i}} \int_{-1}^{1} f(x) e^{i\omega(x-y)^2} dx$$

(where $\omega > 0$) is linear, bounded, of trace class and compact and is known as the Fox–Li operator; its spectrum, $\sigma(\mathcal{F}_\omega)$, is nonempty, discrete and consists of its eigenvalues which are either finite or infinite in number, with at most one point of accumulation at zero [1].

We seek a better theoretical understanding of the behaviour of a generic element of $\sigma(\mathcal{F}_\omega)$ in the asymptotic regime $\omega \gg 1$ as well as a characterisation of its eigenfunctions.

2 Why is the problem important and interesting

A detailed analysis of the spectrum of the Fox–Li operator as well as of its eigenfunctions is particularly relevant to the science of laser beams and resonators whose phenomena is frequently described by linear integral operators defined by kernels that are complex-symmetric but not Hermitian, i.e.,

$$(\mathcal{K}f)(y) := \int_a^b f(x)k(x,y)dx \qquad k(x,y) = k(y,x) \qquad k(x,y) \neq \overline{k(y,x)}$$

of which the Fox–Li operator may be considered as representative, which makes any progress sure to have broad and important applications [2, Section 60].

Also making it a very appealing theme for research, are the facts that it is (mildly) non-normal and falls outside of standard spectral theory [1, 2, 3, 4, 5, 6, 7, 8].

3 State of the art

3.1 Numerical insight

Numerical results available for the eigenvalues and eigenfunctions of the Fox–Li operator enable an important sanity check to any theoretical investigation. These indicate that i) the eigenvalues lie down in a spiral,

contained in the open unit disk, starting from a vicinity of one and rotating clockwise to the origin, which migrates outwards to one and towards the unit circle as ω increases, and that *ii*) the eigenfunctions corresponding to eigenvalues of larger modulus behave as sine and cosine with small perturbations and that those corresponding to eigenvalues of smaller modulus behave as wave packets [3, 9].

3.2 Theoretical developments

Over the years, research on the spectrum of Fox–Li has resulted in a series of papers concerning the spectrum of different, complex-symmetric but not Hermitian, integral operators [3, 4, 8], but despite some theoretical developments regarding the pseudospectrum and the singular values of Fox–Li [5, 6], its spectrum remains a challenge with only a recent clue as to what the Fox–Li spiral might be [7, Theorem 1.2].

4 Our takes on the problem

The talk will concern two of our different approaches, which relate to two different ways of approximating the eigenfunctions: either in inverse powers of large oscillation or in Fourier basis; and conclude with current thoughts on the problem.

References

[1] Albrecht Böttcher, Sergei Grudsky, and Arieh Iserles. The Fox–Li operator as a test and a spur for Wiener-Hopf theory. In Panos M. Pardalos and Themistocles M. Rassias, editors, *Essays in Mathematics and its Applications: In Honor of Stephen Smale's 80th Birthday*. Springer, 2012.

[2] Lloyd N. Trefethen and Mark Embree. *Spectra and Pseudospectra: The Behavior of Nonnormal Matrices and Operators*. Princeton University Press, 2005.

[3] James Alan Cochran and Erold W. Hinds. Eigensystems associated with the complex-symmetric kernels of laser theory. *SIAM Journal on Applied Mathematics*, 26(4):776–786, 1974.

[4] Hermann Brunner, Arieh Iserles, and Syvert P. Nørsett. The spectral problem for a class of highly oscillatory Fredholm integral operators. *IMA Journal of Numerical Analysis*, 30(1):108–130, 2010.

[5] H. J. Landau. The notion of approximate eigenvalues applied to an integral equation of laser theory. *Quarterly of Applied Mathematics*, 35:165–172, 1977.

[6] Albrecht Böttcher, Hermann Brunner, Arieh Iserles, and Syvert P. Nørsett. On the singular values and eigenvalues of the Fox–Li and related operators. *New York Journal of Mathematics*, 16:539–561, 2010.

[7] Albrecht Böttcher, Sergei Grudsky, Daan Huybrechs, and Arieh Iserles. First-order trace formulae for the iterates of the Fox–Li operator. In Harry Dym, Marinus A. Kaashoek, Peter Lancaster, Heinz Langer, and Leonid Lerer, editors, *A Panorama of Modern Operator Theory and Related Topics*, volume 218 of *Operator Theory: Advances and Applications*, pages 207–224. Springer, 2012.

[8] Albrecht Böttcher, Sergei Grudsky, and Arieh Iserles. Spectral theory of large Wiener–Hopf operators with complex-symmetric kernels and rational symbols. *Mathematical Proceedings of the Cambridge Philosophical Society*, 151(1):161–191, 2011.

[9] Hermann Brunner, Arieh Iserles, and Syvert P. Nørsett. The computation of the spectra of highly oscillatory Fredholm integral operators. *Journal of Integral Equations and Applications*, 23(4):467–519, 2011.

On the Collapse of Small Data Self-Gravitating Massless Collisionless Matter

Martin Taylor
Cambridge Centre for Analysis

General relativity is governed by the Einstein equations,

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = T_{\mu\nu}, \tag{1}$$

published by Einstein in 1915 after a seven year struggle to incorporate gravity into his special theory of relativity. This is a system of nonlinear partial differential equations whose unknown is a Lorentz metric g, or, strictly speaking, a Lorentz metric together with a manifold M. Such a *Lorentz manifold* (M, g) solving (1) is called a *spacetime*. Here $R_{\mu\nu}$ denotes the Ricci curvature of g, R the scalar curvature, and $T_{\mu\nu}$ the *stress energy momentum tensor* of the matter present. The Einstein equations therefore relate the curvature of spacetime to the matter present. The equations themselves do not form a closed system, but must be coupled to appropriate equations satisfied by a collection of matter fields defined on M.

In the vacuum case, where $T_{\mu\nu} = 0$, (1) reduce to the *vacuum Einstein equations*,

$$R_{\mu\nu} = 0. \tag{2}$$

What's quite remarkable, and in stark contrast to the preceeding Newtonian theory of gravity, is that the Einstein equations are non-trivial even in the absence of matter. In fact, there are solutions of (2) which contain regions "unable to send signals to far away observers". Such solutions are known as *black hole spacetimes*.

The trivial solution of the vacuum Einstein equations (2) is Minkowski spacetime, which is $\mathbb{R}^{3+1}$ together with the Minkowski metric $m = -dt^2 + (dx^1)^2 + (dx^2)^2 + (dx^3)^2$. This is the Lorentzian analogue of Euclidean space; not only does the Ricci curvature of this spacetime vanish, but the full Riemann tensor in fact vanishes. In particular, it is geodesically complete and contains no such black hole regions.

Two of the most fundamental aspects of the Einstein equations are their *general covariance* and their *hyperbolicity*. Their general covariance, or in more modern language their *geometric structure*, in fact obscures the fact they are hyperbolic. This is of vast importance however as it means the equations have a well posed Cauchy problem and hence the dynamic viewpoint is the correct one. One can therefore talk about the dynamic stability of particular of particular solutions. This talk will focus on the stability of Minkowski space. The most direct way to see the hyperbolicity is via a specific coordinate choice, called the *harmonic gauge*. In such coordinates, the Einstein equations reduce to a system of quasilinear wave equations. Despite much work on the well posedness of the Cauchy problem for equations of this form in the early half of the twentieth century, it wasn't until the seminal 1952 work of Choquet-Bruhat [2] that the question of local existence for the full vacuum Einstein equations (2) was resolved. The geometric initial data for (2) consists of a triple $(\Sigma, \bar{g}, K)$, where Σ is to be a spacelike hypersurface in the resulting spacetime (M, g), and $\bar{g}, K$ are to be the induced first and second fundamental forms respectively. Whereas initial data for systems of quasilinear wave equations can be freely specified, this is manifestly not the case for (2) as if $(\Sigma, \bar{g}, K)$ indeed imbedded into a Ricci flat manifold (M, g), the contracted Gauss and Codazzi equations would imply

$$\bar{R} + (\mathrm{tr}K)^2 - |K|^2_{\bar{g}} = 0, \qquad \overline{\mathrm{div}} K - d\mathrm{tr}K = 0, \tag{3}$$

where $\bar{R}$ is the scalar curvature of $(\Sigma, \bar{g})$.

Theorem 1 (Choquet-Bruhat [2]). *Given a smooth vacuum initial data set* $(\Sigma, \bar{g}, K)$ *satisfying the Einstein constraint equations* (3)*, there exists a smooth development of the data, i.e., a Lorentz manifold* (M, g) *satisfying* (2) *and a smooth embedding* $i : \Sigma \to M$ *such that* $i_*\bar{g}, i_*K$ *are the induced first and second fundamental forms of* $i(\Sigma)$ *respectively.*

In 1969, Choquet-Bruhat–Geroch [3] showed that there exists a unique developent of initial data satisfying an appropriate maximality condition. Given Theorem 1, one can ask the following stability question: *what happens when one evolves data for the Einstein vacuum equations which is suitably "close" to initial data corresponding to Minkowski spacetime?* This question was answered, providing the most celebrated global result on the Einstein equations, in the monumental 1993 work of Christodoulou–Klainerman.

Theorem 2 (Christodoulou–Klainerman [4]). *If* $(\Sigma, \bar{g}, K)$ *is an initial data set for the vacuum Einstein equations satisfying an appropriate asymptotic flatness condition, and is sufficiently close to Minkowski initial data, then the resulting maximal development is geodesically complete and the spacetime approaches Minkowski spacetime (with quantitative decay rates) in every direction. Moreover, a complete future null infinity* $\mathcal{I}^+$ *can be attached to the spacetime such that* $J^-(\mathcal{I}^+) = M$.

The last sentence is a precise way of saying that the spacetime has no black hole regions or naked singularities. A natural question to ask now is, *does a similar thing happen in the presence of matter?* One of the simplest matter models is collisionless matter. The evolution of collisionless matter in general relativity is described by the Einstein–Vlasov system, which is (1) coupled with the Vlasov equation

$$T^{\mu\nu} = \int_{\pi^{-1}(x)} p^\mu p^\nu f, \qquad p^\alpha \partial_{x^\alpha} f - \Gamma^\alpha_{\beta\gamma} p^\beta p^\gamma \partial_{p^\alpha} f = 0.$$

Here (p^α, x^α) are coordinates on the tangent bundle TM, $\pi : TM \to M$ is the natural projection, and the unknown is a density function $f : P \to [0, \infty)$, where $P \subset TM$ is a subset of the causal vectors. Given an initial data set which satisfies an appropriate analogue of (3), one can prove a version of Theorem 1 for the Einstein–Vlasov system [1]. The study of the stability of this system was initiated by Rein and Rendall [6], who looked at the above question in spherical symmetry for the case where all of the particles have the same positive mass. The massless case was resolved in spherical symmetry by Dafermos.

Theorem 3 (Dafermos [5]). *Given regular spherically symmetric asymptotically flat initial data for the Einstein–Vlasov system where* f_0 *is compactly supported on the set of null vectors, if the data is sufficiently close to trivial Minkowski data, then the resulting spacetime is geodesically complete and the components of the energy momentum tensor with respect to a naturally defined null coordinate system decay.*

The talk will discuss my recent efforts to remove the assumption of spherical symmetry from Theorem 3.

References

[1] Y. Choquet-Bruhat *Problème de Cauchy pour le système intégro-différentiel d'Einstein-Liouville* Ann. Inst. Fourier, **21** (1971) 181-201.

[2] Y. Choquet-Bruhat *Théreme d'existence pour certains systèmes déquations aux dérivées partielles non linéaires* Acta Math., **88** (1952) 141-225.

[3] Y. Choquet-Bruhat and R. Geroch *Global Aspects of the Cauchy Problem in General Relativity* Comm. Math. Phys., **14** (1969) 329-335.

[4] D. Christodoulou and S. Klainerman, *The Global Nonlinear Stability of the Minkowski Space* Princeton Mathematical Series, Vol. 41 (Princeton University Press, 1993).

[5] M. Dafermos *A Note on the Collapse of Small Data Self-Gravitating Massless Collisionless Matter* J. Hyperbol. Differ. Equations, **3** (2006) 589-598.

[6] G. Rein and A. D. Rendall, *Global Existence of Solutions of the Spherically Symmetric Vlasov–Einstein System with Small Initial Data*, Commun. Math. Phys. **150** (1992) 561-583.

Electrokinetic instbility in the Poiseuille flow

Lukáš Vermach
Cambridge Centre for Analysis

The talk is concerned with a flow of electrically charged liquid running through a narrow channel with charged walls. Such a problem belongs to a wider set of effects known as the electrokinetic phenomena. Hunter [1] distinguish four kinetic phenomena: Electroosmosis effect, Streaming potential production, Electrophoresis effect and Sedimentation potential. In the case of the pressure-driven flow through a channel the effect of streaming potential is relevant.

As the list of the effects suggests, the phenomena occur in many natural processes and industrial applications. For instance, the topic is important in the context of the petroleum industry and aircraft fuel tanks design. The presence of the electric charge in fuel reservoirs and related distributive nets represents a serious safety danger since the high concentration of electric charge can initiate spontaneous ignition in the greatly flammable gasses above the fuel. Thus, understanding the evolution (particularly, the rate of decay) of the electric charge carried by a fluid is of a great interest with respect to assessing electrostatic hazards in fuel tanks.

All the electrokinetic effects are concerned with coupling between mechanical and electrical behaviour in fluids which is due to the formation of the electric double layer (EDL) on the solid-liquid interface. The EDL comprises two layers, each being formed by ions of opposite sign. One is stationary, directly attached to the solid surface, the other belongs to the fluid. As the charge carriers within the liquid are drifted downstream along the channel an axial streaming current is generated. Accumulation of the charged particles in the downstream region produces an additional electric field which in turn results in the electric current in the opposite (upstream) direction. This process then naturally affects the characteristics of the flow.

The state of Newtonian, incompressible fluid carrying charged particles (for simplicity the symmetric electrolyte is assumed) is described by velocity $\mathbf{u}$, pressure p, ionic concentrations per unit volume n_+, n_- and electrostatic potential φ. The evolution of these quantities is governed by Navier Stokes equations, continuity equation, electrostatic Maxwell equations and charge conservation (NernstPlanck equation):

$$\rho\left(\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u}\cdot\nabla\mathbf{u}\right) = -\nabla p - ez(n_+ - n_-)\nabla\varphi + \mu\Delta\mathbf{u}, \qquad \nabla\cdot\mathbf{u} = 0,$$

$$-\epsilon\Delta\varphi = ez(n_+ - n_-), \qquad \nabla\mathbf{i}_\pm \pm e\frac{\partial n_\pm}{\partial t} = 0,$$

where $\mathbf{i}_\pm$ denotes the electric current density of corresponding ionic types. The symbols ρ, e, μ and ϵ denote fluid density, elementary charge, dynamic viscosity and permittivity respectively; z stands for the valence number. The electric current density is given

$$\mathbf{i}_\pm = -ezD\nabla n_\pm \mp n_\pm e^2 z^2 D\frac{\nabla\varphi}{k_B T} + ezn_\pm\mathbf{u},$$

where D denotes the ionic diffusion coefficient, k_B is the Bolzmann constant and T is the thermodynamic temperature.

We start off the discussion by analysis of the steady state solution for pressure driven channel flow. In particular, the research is focused on the influence of the streaming potential on the underlying parabolic velocity profile. It turns out that if the electrokinetic phenomenon is strong enough, the reverse flow in the boundary region of the channel can occur. Such an effect is caused by the conduction current due to the streaming field established as a consequence of the downstream migration of the charge carriers in the centre region of the channel. The counter motion (with respect to the pressure-driven flow) is realised in boundary region as the concentration of one ionic type is amplified here due to the EDL (as opposed to the centre of the channel where the net charge is zero).

Finally, following the [2] and [3], the stability of the steady state solution is examined. It is well known that for pure pressure-driven flow with absence of electric charge the stability breaks down at certain critical Reynolds number Re_c. The modal theory predicts $\mathrm{Re}_c \approx 5772$ (Orszag 1971, [4]). We explore the influence of the electric charge on this result. Consequently, we address the question whether the presence of electric charge changes the position of the most unstable mode and we study the sensitivity to the change of the dimensionless parameters controlling the flow profile.

References

[1] Hunter RJ. 1981. *Zeta Potential in Colloid Science: Principles and Applications.* Academic Press

[2] Schmid PJ, Henningson DS. 2001. *Stability and Transition in Shear Flows.* New York: Springer-Verlag

[3] Schmid PJ. 2007. *Nonmodal Stability Theory.* Annu. Rev. Fluid Mech. 39:129-62.

[4] Orszag SA. 1971. *Accurate solution of the Orr-Sommerfeld stability equation.* J. Fluid. Mech. 29:93-111.