UNIVERSITY OF CALIFORNIA
Los Angeles

Optimization of Signal Segmentation, Signal Recovery, and Limited Current

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Mathematics

by

Hung-Hsu Chou

2019

ABSTRACT OF DISSERTATION

Optimization of Signal Segmentation, Signal Recovery, and Limited Current

by

Hung-Hsu Chou

Doctor of Philosophy in Mathematics

University of California, Los Angeles, 2019

Professor Russel E. Caflisch, Chair

The first chapter is based on applying the Poisson summation formula to a constrained optimization problem. Motivated by Shannon sampling theorem and results on shift-invariant subspaces, we establish a compatible framework for the two key factors: the accuracy constraint, which is described in the frequency space, and the efficiency function, which is expressed in the regular space. We derive the optimal wavelet, denoted as the double-sinc function, that obtains the smallest support while remaining first order accuracy. Based on this wavelet, we further improve its accuracy by loosen up the constraint in support and manage to achieve nearly optimal efficiency.

The goal of the second chapter is to recover the underlying signal from its superposed randomly-shifted noisy measurement, motivated by multi-reference alignment and Cryo-EM problem. The general setting is that we observe samples from noisy signal that is acted by a random group action, and we would like eliminate those noises, one type at a time. In our particular setting, rational Fourier monomials and total Fourier product are invariants under the group action and hence partially remove the effect of the noise from the group action. We then apply central limit theorem to eliminate the Gaussian noise. Finally, we apply the split Bregman algorithm in compressed sensing to obtain an explicit solution assuming that the signal is compactly supported.

The third chapter is dedicated to applying a variation principle to the Euler-Poisson equation for periodic flow in a diode to optimize the flux difference, which could potentially exceed the Child-Langmuir limit. We derive a set of dual equations and boundary conditions and use upwind method to solve both the forward and backward equations. In our numerical experiment we derive a periodic solution whose flux goes above the CL limit before the physical setting or the method of characteristics breaks down.

The dissertation of Hung-Hsu Chou is approved.

Inwon Kim

Stanley J. Osher

Monica Visan

Russel E. Caflisch, Committee Chair

University of California, Los Angeles

2019

To the world and myself

Contents

List of Figures

ACKNOWLEDGEMENTS

I am grateful to have Russel Caflisch as my advisor, who has been providing me insightful guidance. I would like to thank Afonso Bandeira and Inwon Kim for their reading groups and advice, as well as all my committee members. I would also like to thank my colleague Jonathan Siegel for the enormous amount of useful discussions. My research was conducted under the support of UC Regents, NSF, and AFOSR grant.

VITA

Education

2013	B.S. in Physics
	College of Creative Studies, UCSB
2017	Adjunct Instructor
	Courant Institute of Mathematical Sciences, NYU

Publication

- Russel E. Caflisch, Hung Hsu Chou, and Jonathan W. Siegel (2018). **Accuracy, Efficiency and Optimization of Signal Fragmentation.** submitted to *Multiscale Modeling and Simulation*

- Russel E. Caflisch, Hung Hsu Chou, and Jonathan W. Siegel (2018). **Efficient Frame Projection of Amplitude-Modulated Signals.** *UCLA CAM report*

- Michael D. Johnson, Hung Hsu Chou, Carl R. Gwinn (2013). **Optimal Correlation Estimators for Quantized Signals.** *The American Astronomical Society*, 752(2), 135-142. arXiv:1210.7271

Chapter 1

Signal Segmentation of Amplitude-Modulated Signal

Signals are often analyzed in the frequency space despite the fact that our devices operate in regular space. Bridges between those spaces have received much attention and led to great results, the Poisson summation formula and Shannon sampling theorem for instance. Our work focuses on the finding an accurate and efficient way to segment amplitude-modulated signals, with applications to antenna transmission.

Consider a signal $f \in L^2(\mathbb{R})$ whose Fourier transform is supported on a small neighborhood of radius ϵ around ± 1. Signals of this type are called amplitude-modulated (AM) signal with bandwidth ϵ. Our first goal is to find wavelets $\phi \in L^2_c(\mathbb{R})$ such that the approximation

$$f(t) \approx \sum_{n \in \mathbb{Z}} f(n\Delta)\phi(t - n\Delta) \tag{1.1}$$

is highly accurate for AM signals, where $N \in \mathbb{N}$, and $\Delta = \frac{2\pi}{N}$ represents the spacing between wavelets. A more compact expression of (1.1) is

$$f \approx \phi * C_\Delta \cdot f \tag{1.2}$$

where $C_\Delta := \sum_{n\in\mathbb{Z}} \delta_{n\Delta}$ is the sum of Dirac distributions over the lattice $\Delta\mathbb{Z}$, also known as the Dirac comb.

The second goal is to minimize the support size of wavelets. In antenna transmission, the support size of wavelets correspond to the amount of time each device needs to be turned on. The signal is fully delivered only after all device finish their tasks, and hence it is undesirable to have wavelets with large support. For instance, in Shannon sampling theorem, the wavelets are not compactly supported and requires modification in practice [1]. Errors from those modification are often complicated if not intractable, and hence we want to avoid such by requiring our wavelets to be compactly supported.

The third goal is to maximize the efficiency of our approximation. Inefficiency are mostly due to cancellation between wavelets, and can be costly in practice. Our formulation of efficiency is based on the energy in the far-field energy, the Ohmic heating, and the radiation [2].

In summary, we would like to find wavelets ϕ such that for any AM signal f,

1. the projection error $||f - \phi * C_\Delta \cdot f||_2$ is small,
2. the support size $|\Omega_\phi|$ is small,
3. the efficiency $E(\phi) := \frac{\hat{\phi}(1)^2}{|\Omega_\phi|\cdot||\phi||_2^2}$ is large.

Notation-wise, denote $\hat{f}(k)$ or $\mathcal{F}(f)(k) := \int_{-\infty}^{\infty} f(t)e^{-ikt}dt$ to be the Fourier transform of f evaluated at k, and Ω_ϕ to be the support of ϕ.

We first provide motivation of our formulation in section 1.1 and formally state the accuracy constraint and the scaling of the error in projection in section 1.2. In section 1.3 we derive the lower bound of support sizes subject to the accuracy constraint. In section 1.4 we formulate the efficiency optimization problem in two ways, one through analytic calculation and one through numerical simulation. In section 1.5 we compute the efficiency of the wavelets we derive and compare them with other typical wavelets.

1.1 Motivation from Poisson and Shannon

In signal processing, signals are often easier to describe in the frequency space. By Parseval's identity, the projection error can be expressed as

$$\begin{aligned} ||f - \phi * C_\Delta \cdot f||_2 &= ||\mathcal{F}(f - \phi * C_\Delta \cdot f)||_2 \\ &= ||\hat{f} - \hat{\phi} \cdot NC_N * \hat{f}||_2 \end{aligned}$$

due to the Poisson summation formula (A.2). Therefore it suffices to find ϕ such that $||\hat{f} - \hat{\phi} \cdot NC_N * \hat{f}||_2$ is small. For simplicity, we will absorb the factor N into the wavelet ϕ.

An application of the Poisson summation formula is the Shannon sampling theorem, which plays a important role in signal processing. The techniques in the Shannon sampling theorem (1.1) is simple yet inspiring, which will be shown as we reconstructed the proof of theorem 1.1.

Theorem 1.1. *(Shannon sampling theorem[3])*
Fix $N > 0$ and $\Delta = \frac{2\pi}{N}$. A signal $f \in L^2(\mathbb{R})$ whose Fourier transform is supported on $(-N/2, N/2)$ can be fully reconstructed from its samples on the lattice $\Delta\mathbb{Z}$. The reconstruction takes the form

$$f = \phi_\Delta * C_\Delta \cdot f \tag{1.3}$$

where $\phi_\Delta(t) := \sin(\Delta^{-1}t)/(\Delta^{-1}t)$ is called the sinc function.

Proof. The core idea is to periodize $\hat{f}$ with C_N and unperiodized it with $\hat{\phi}$, illustrated by figure 1.1.

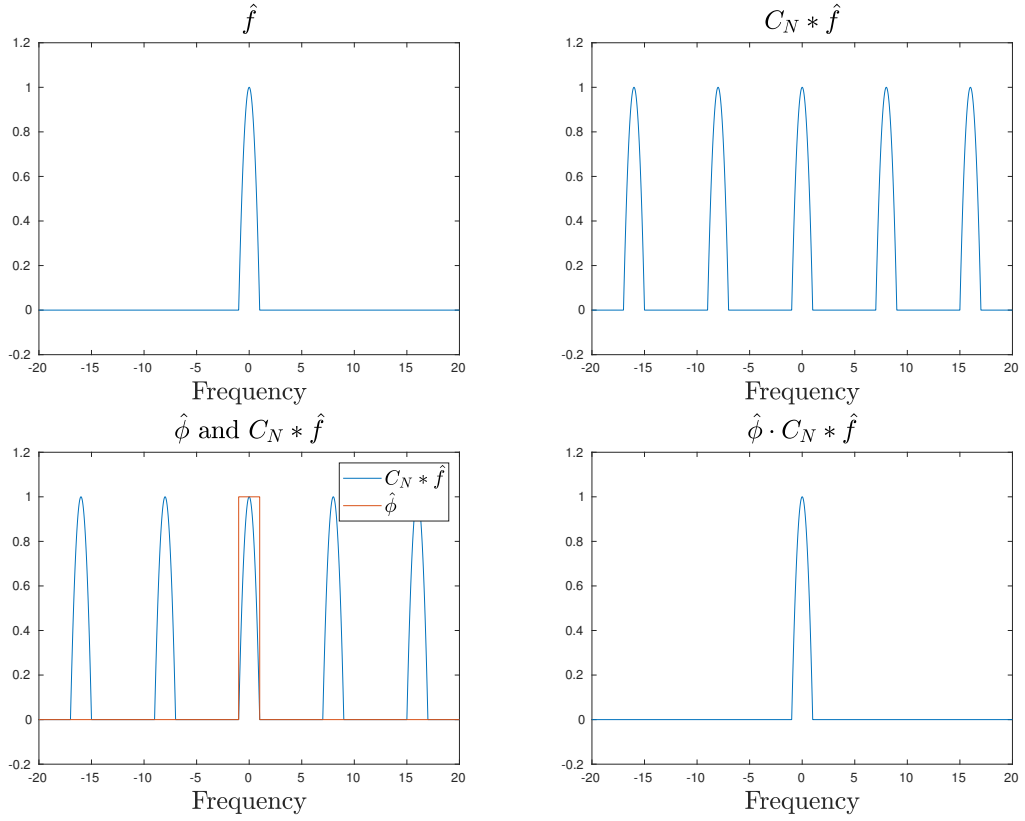

FIGURE 1.1: Periodization and unperiodization of band-limited signals in frequency space, with perfect reconstruction

Clearly, if $\hat{\phi}$ is the indicator function of $(-N, N)$, then the reconstruction is perfect and $\hat{f} = \hat{\phi} \cdot C_N * \hat{f}$. Note $\hat{\phi} = N\hat{\phi}_\Delta$, and hence f can be expressed as

$$f = \mathcal{F}^{-1}\hat{f} = \mathcal{F}^{-1}\left(N\hat{\phi_\Delta} \cdot C_N * \hat{f}\right) = \phi_\Delta * C_\Delta \cdot f$$

where $\phi_\Delta(t) := \sin(\Delta^{-1}t)/(\Delta^{-1}t)$. □

We would like to apply similar techniques to AM signals, but with the constraint that our wavelet must be compactly supported in the regular space, which means it can not be compactly supported in the frequency space. In exchange, we will sacrifice some accuracy and the error scales with ϵN^{-1}; in the limit as ϵN^{-1} goes to zero we will obtain perfect reconstruction of am signals. Our method, similar to the one's used by Shannon, is illustrated in figure 1.2.

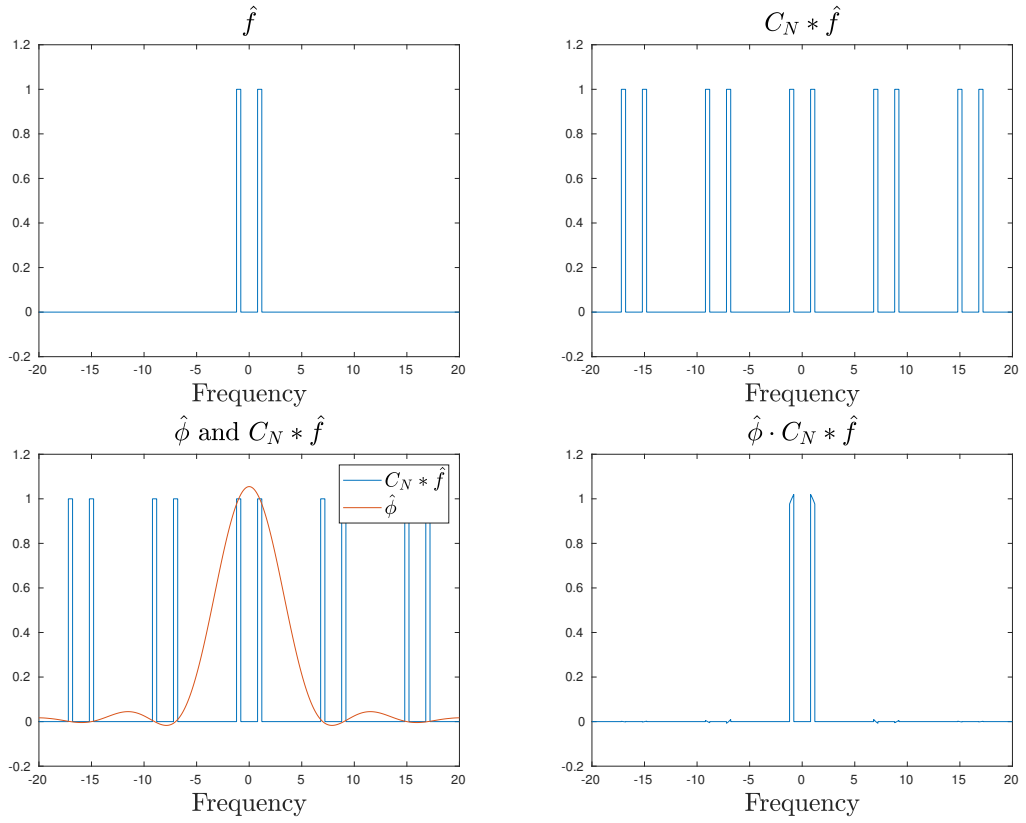

FIGURE 1.2: Periodization and unperiodization of AM signals in frequency space, with imperfect reconstruction

1.2 Projection Accuracy

Projections to partial wavelet spaces have received much attention in signal processing literature [4] due to its applications. Along with results from frame theory [5] and complex analysis [6], it was shown that the accuracy of such projections are related to the order of zeros on some lattice in the frequency space in the following sense:

Definition 1.2. Let $\epsilon > 0$, the space of ϵ-band-limited functions H_ϵ is given by

$$H_\epsilon := \{f \in L^2 \ s.t. \ \hat{f}(k) = 0 \text{ if } |k| \geq \epsilon\}. \tag{1.4}$$

Let $\phi \in L^2(\mathbb{R})$ with $\|\phi\|_2 = 1$. The partial wavelet space of ϕ, $S(\phi)$, is defined by

$$S(\phi) := \overline{\text{span}\{\phi(x-n)\}_{n\in\mathbb{Z}}}. \tag{1.5}$$

Let $\phi \in L^2(\mathbb{R})$ and $\epsilon > 0$. The accuracy of ϕ at bandwidth ϵ is

$$A_\phi(\epsilon) := \sup_{f\in H_\epsilon,\ \|f\|_2=1} \|P_{S(\phi)}f - f\|_2^2 \tag{1.6}$$

where P stands for projection.

Theorem 1.3 ([6]). *Assume that $\phi \in L^2(\mathbb{R})$ has compact support and*

$$R_\phi(k) := \sum_{n\in\mathbb{Z}} |\hat{\phi}(k-n)|^2 \tag{1.7}$$

is bounded and bounded away from 0. Let $L > 0$ be an integer. Then

$$A_\phi(\epsilon) = O(\epsilon^{2L}) \tag{1.8}$$

iff $\mathcal{F}(\phi)^{(l)}(n) = 0$ for all $n \in \mathbb{Z}\backslash\{0\}$ and $l = 0, ..., L-1$, i.e. if and only if the Fourier transform of ϕ vanishes to order L at all non-zero integers.

Theorem 1.3 describes the conditions for accurate projections and their corresponding error. Though larger L leads to smaller error, for simplicity we will consider the simplest case where $L = 1$. Since AM signals have Fourier transform supported around ± 1 and the spacing between our wavelets are Δ rather than 1, we are interested in the modified error

$$\tilde{A}_\phi(\epsilon) := \sup_{f\in H_\epsilon \pm 1,\ \|f\|_2=1} \|P_{S_\Delta(\phi)}f - f\|_2^2 \tag{1.9}$$

where $S_\Delta(\phi) := \overline{\text{span}\{\phi(x-n\Delta)\}_{n\in\mathbb{Z}}}$. A natural condition for wavelets in AM signals, motivated by figure 1.2 and theorem 1.3, is defined in definition 1.4.

Definition 1.4. Fix $N \geq 3$. We say $\phi \in L_c^2(\mathbb{R})$ is first order accurate at the scale of N for AM signals if

$$\hat{\phi}(\pm 1 + N\mathbb{Z}\backslash\{0\}) = 0 \ , \ \hat{\phi}(\pm 1) = 1. \tag{1.10}$$

We will refer (1.10) as the accuracy constraint.

Theorem 1.5. *Fix $N \geq 3$ and $\Delta = \frac{2\pi}{N}$. Suppose ϕ is first order accurate at the scale of N for AM signals, and f is an AM signal with bandwidth ϵ. The projection error scales like*

$$||f - \phi * C_\Delta \cdot f||_2 = O(\epsilon\Delta). \tag{1.11}$$

The proof of theorem 1.5 is very similar to the proof of theorem 1.3. As a result of theorem 1.5, sinusoidal signals can be perfectly reconstructed by first order accurate wavelets.

The accuracy constraints involve infinitely many qualities, which can be hard to work with. Fortunately, it can be translated into a system of two equations in the regular space according theorem 1.6, which will be useful in the later sections.

Theorem 1.6. *Fix $N \geq 3$ and $\Delta = \frac{2\pi}{N}$. Suppose ϕ is a Schwartz function in $\mathbb{R}$. Then ϕ is first order accurate at the scale of N for AM signals if and only if*

$$U_\phi^\pm := N^{-1} C_\Delta * Q_\phi^\pm = 1 \tag{1.12}$$

where $Q_\phi^\pm := e^{\mp it}\phi(t)$.

Proof. By the Poisson summation formula,

$$\mathcal{F}(U_\phi^\pm) = \mathcal{F}\left(\frac{1}{N} C_\Delta * e^{\mp it}\phi(t)\right) = C_N \cdot \delta_{\pm 1} * \mathcal{F}(\phi). \tag{1.13}$$

Hence ϕ is first order accurate at the scale of N for AM signals if and only if $\mathcal{F}(U_\phi^\pm) = \delta_0$, which is equivalent to $U_\phi^\pm = 1$. □

1.3 Support Bound

Since the accurate wavelets for AM signals are not required to be compactly supported in the frequency space, we expect the existence of accurate compactly supported wavelets. While the existence of such wavelets are easy to show, we are further interested in finding the ones that have the minimal support, if there a minimum do exist. Using the equivalent accuracy constraints 1.12, we manage to derive the minimum support and its corresponding minimizer, as described in theorem 1.7.

Theorem 1.7. *Fix $N \geq 3$ and $\Delta = \frac{2\pi}{N}$. If ϕ is first order accurate at the scale of N for AM signals, then $|\Omega_\phi| \geq 2\Delta$. If $|\Omega_\phi| = 2\Delta$, the ϕ equals to the double sinc function, denoted as ψ and defined as*

$$\psi := \begin{cases} \frac{N}{\sin(\Delta)} \sin(\Delta - |t|) & \text{if } |t| \leq \Delta \\ 0 & \text{otherwise} \end{cases} \tag{1.14}$$

Proof. Suppose $\phi \in D_{1,N} \cap D_{-1,N}$ with support $\Omega_\phi \subset [-\Delta, \Delta]$. Decompose ϕ into $\phi_1(t) = \phi(t - \Delta)\chi_{[0,\Delta]}(t)$ and $\phi_2(t) = \phi(t)\chi_{[0,\Delta]}(t)$ (so that $\phi = \phi_1(t + \Delta) + \phi_2(t)$), where χ are indicator functions. By theorem 1.6 we have

$$N\chi_{[0,\Delta]}(t) = e^{-it}(e^{i\Delta}\phi_1 + \phi_2) \tag{1.15}$$

$$N\chi_{[0,\Delta]}(t) = e^{it}(e^{-i\Delta}\phi_1 + \phi_2) \tag{1.16}$$

The solution to the system is

$$\phi_1(t) = \frac{N}{\sin(\Delta)} \sin(t)\chi_{[0,\Delta]}(t) \tag{1.17}$$

$$\phi_2(t) = \frac{N}{\sin(\Delta)} \sin(\Delta - t)\chi_{[0,\Delta]}(t) \tag{1.18}$$

where we recover (1.14). □

According to theorem 1.6, the projection error of ψ should go to zero when f is sinusoidal, which implies the following corollary.

Corollary 1.8. *Sinusoidal signals can be perfectly reconstructed by the linear combination of shifted double sinc function ψ defined in (1.14). In particular, for integer $N \geq 3$,*

$$\sin(t) = \frac{1}{N} \sum_{n=-\infty}^{\infty} \sin(n\Delta)\psi(t - n\Delta) \tag{1.19}$$

where $\Delta = \frac{2\pi}{N}$.

Proof. See A.2. □

1.4 Efficiency Optimization

Under the assumption of perfect superposition, i.e. the far-field signal is equal to the sum of the derivatives of the currents in antennas, efficiency can be defined as the ratio between energy density in the far field and the energy dissipation in the antennae, which is due to radioactive loss and Ohmic heating. Suppose the efficiency is defined as

$$E_{\text{eff}} := \frac{E_{\text{far}}}{E_{\text{rad}} + E_{\text{ohm}}}. \tag{1.20}$$

In the limit when the number wavelets N is large and the number of antenna M is comparable to N, [2] states that the efficiency for such projection is approximately

$$E_{\text{eff}} \approx E_1(\phi) := \frac{\hat{\phi}(1)^2}{|\Omega_\phi| \cdot ||\phi||_2^2} \tag{1.21}$$

$$\approx E_2(\phi) := \frac{||\phi||_1^2}{|\Omega_\phi| \cdot ||\phi||_2^2} \tag{1.22}$$

after normalization. Roughly speaking, the numerator is the far field energy and denominator is the radioactive loss; the Ohmic heating is less significant in the regime where M and N are large. Here we list two efficiency formulations (that are approximately the

same when $M, N \ll 1$) because (1.21) is easier to work with analytically, as shown in section 1.4.1 while (1.22) is easier to work with numerically, as shown in section 1.4.3.

Our goal at this stage is to maximize the efficiency, while maintaining accuracy and small support. The accuracy constraints in theorem 1.6 implies that $|\Omega_\phi| \geq 2\Delta$, due to corollary 1.7. To add some degrees of freedom in our optimization problem, we define our objective function to be

$$\max_{\phi} E_{\text{eff}}(\phi) \quad \text{subject to} \quad U_\phi^{\pm} = 1 \text{ and } |\Omega_\phi| \leq 4\Delta. \tag{1.23}$$

1.4.1 Derivation of Analytic Optimizer

Consider

$$\max_{\phi} E_1(\phi) \quad \text{subject to} \quad U_\phi^{\pm} = 1 \text{ and } |\Omega_\phi| \leq 4\Delta. \tag{1.24}$$

In this section we derive an analytic solution to problem 1.24. Since the conditions $U_\phi^{\pm} = 1$ implies that $\hat{\phi}(1) = 1$, by fixing the upper bound of support size, the problem 1.24 can be reduced to

$$\min_{\phi} ||\phi||_2^2 \quad \text{subject to} \quad U_\phi^{\pm} = 1 \text{ and } |\Omega_\phi| \leq 4\Delta. \tag{1.25}$$

Define $\phi_n := \phi(t + (n-3)\Delta)\chi_{[0,\Delta]}(t)$ for $1 \leq n \leq 4$; equivalently,

$$\phi(t) = \phi_1(t + 2\Delta) + \phi_2(t + \Delta) + \phi_3(t) + \phi_4(t - \Delta)). \tag{1.26}$$

Then the constraints in (1.25) become

$$N\chi_{[0,\Delta]}(t) = Q_{u,p}^{+} = e^{-it}[e^{2i\Delta}\phi_1 + e^{i\Delta}\phi_2 + \phi_3 + e^{-i\Delta}\phi_4] \tag{1.27}$$

$$N\chi_{[0,\Delta]}(t) = Q_{u,p}^{-} = e^{it}[e^{-2i\Delta}\phi_1 + e^{-i\Delta}\phi_2 + \phi_3 + e^{i\Delta}\phi_4]. \tag{1.28}$$

Consider e^{it}(1.27) $\pm\, e^{-it}$(1.28), we get

$$
\begin{aligned}
& N\cos(t) = \cos(2\Delta)\phi_1 + \cos(\Delta)\phi_2 + \phi_3 + \cos(\Delta)\phi_4 \\
& N\sin(t) = \sin(2\Delta)\phi_1 + \sin(\Delta)\phi_2 - \sin(\Delta)\phi_4 \\
\implies & N\begin{bmatrix}\cos(t)\\ \sin(t)\end{bmatrix} = \begin{bmatrix}\cos(2\Delta) & \cos(\Delta)\\ \sin(2\Delta) & -\sin(\Delta)\end{bmatrix}\begin{bmatrix}\phi_1\\ \phi_4\end{bmatrix} + \begin{bmatrix}\cos(\Delta) & 1\\ \sin(\Delta) & 0\end{bmatrix}\begin{bmatrix}\phi_2\\ \phi_3\end{bmatrix} \\
\implies & \begin{bmatrix}\phi_2\\ \phi_3\end{bmatrix} = \frac{N}{\sin(\Delta)}\begin{bmatrix}\sin(t)\\ \sin(\Delta - t)\end{bmatrix} + \begin{bmatrix}-2\cos(\Delta) & 1\\ 1 & -2\cos(\Delta)\end{bmatrix}\begin{bmatrix}\phi_1\\ \phi_4\end{bmatrix}
\end{aligned}
$$

Hence we arrive a simpler form of objective:

$$
\begin{aligned}
\min_{\phi_1,\phi_4} & ||\phi_1||_2^2 + ||\phi_4||_2^2 + \left|\left|\phi_4 - 2\cos(\Delta)\phi_1 + N\sin(\Delta)^{-1}\sin(t)\right|\right|_2^2 \\
& + \left|\left|\phi_1 - 2\cos(\Delta)\phi_4 + N\sin(\Delta)^{-1}\sin(\Delta - t)\right|\right|_2^2
\end{aligned} \tag{1.29}
$$

The symmetry suggests that $\phi_2(t) = \phi_3(\Delta - t)$. Solving the system by differentiating with respect to ϕ_1 and ϕ_4, we get

$$
\begin{aligned}
0 &= (2 + 4\cos^2(\Delta))\phi_1 - 4\cos(\Delta)\phi_4 + \frac{N}{\sin(\Delta)}(\sin(\Delta - t) - 2\cos(\Delta)\sin(t)) \\
0 &= (2 + 4\cos^2(\Delta))\phi_4 - 4\cos(\Delta)\phi_1 + \frac{N}{\sin(\Delta)}(\sin(t) - 2\cos(\Delta)\sin(\Delta - t)) \\
\implies \begin{bmatrix}\phi_1\\ \phi_4\end{bmatrix} &= c\begin{bmatrix}2 + 4\cos^2(\Delta) & 4\cos(\Delta)\\ 4\cos(\Delta) & 2 + 4\cos^2(\Delta)\end{bmatrix}\begin{bmatrix}2\cos(\Delta) & -1\\ -1 & 2\cos(\Delta)\end{bmatrix}\begin{bmatrix}\sin(t)\\ \sin(\Delta - t)\end{bmatrix} \\
&= c\begin{bmatrix}8\cos^3(\Delta) & 4\cos^2(\Delta) - 2\\ 4\cos^2(\Delta) - 2 & 8\cos^3(\Delta)\end{bmatrix}\begin{bmatrix}\sin(x)\\ \sin(\Delta - x)\end{bmatrix}
\end{aligned}
$$

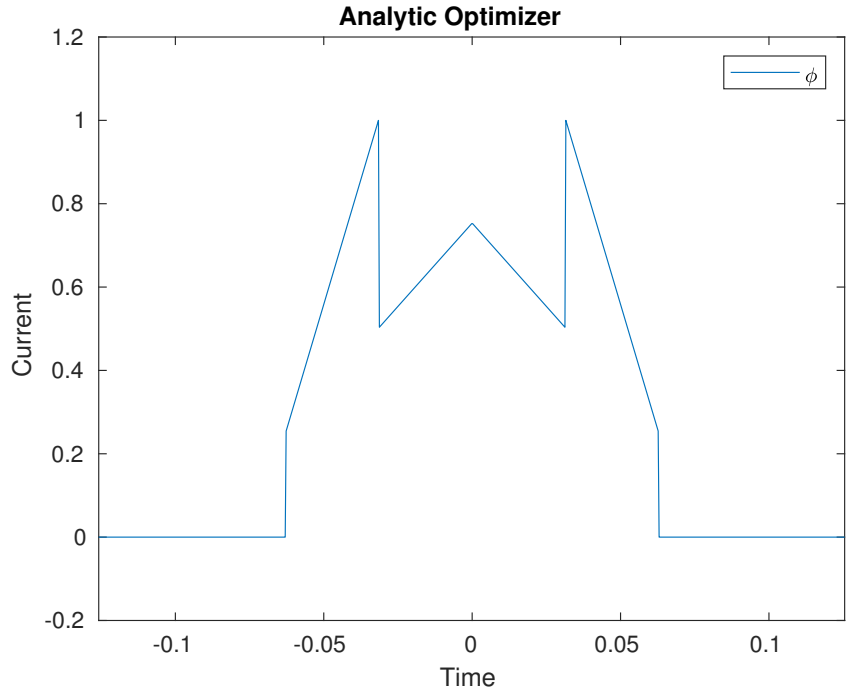

FIGURE 1.3: Efficiency maximizer ϕ

where $c = \frac{N}{4\sin(\Delta)(1+4\cos^4(\Delta))}$. Hence the full solution is

$$\begin{bmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \\ \phi_4 \end{bmatrix} = c \begin{bmatrix} 8\cos^3(\Delta) & 4\cos^2(\Delta) - 2 \\ 4\cos^2(\Delta) + 2 & 4\cos(\Delta) \\ 4\cos(\Delta) & 4\cos^2(\Delta) + 2 \\ 4\cos^2(\Delta) - 2 & 8\cos^3(\Delta) \end{bmatrix} \begin{bmatrix} \sin(x) \\ \sin(\Delta - x) \end{bmatrix} \tag{1.30}$$

shown in figure 1.3 and 1.4. Note that the function ϕ is symmetric since $\phi_2(t) = \phi_3(\Delta - t)$ and $\phi_1(t) = \phi_4(\Delta - t)$. From now on we will denote this analytic optimal wavelet, described in (1.30), as ϕ. We will analyze its properties in section 1.4.2.

1.4.2 Analysis on Analytic Optimizer

Due to constraint (1.10), which is equivalent to (1.12), the double sinc function ψ and the optimal wavelet ϕ share the same roots on the lattice $\pm 1 + \mathbb{Z}\backslash\{0\}$ in the Fourier space.

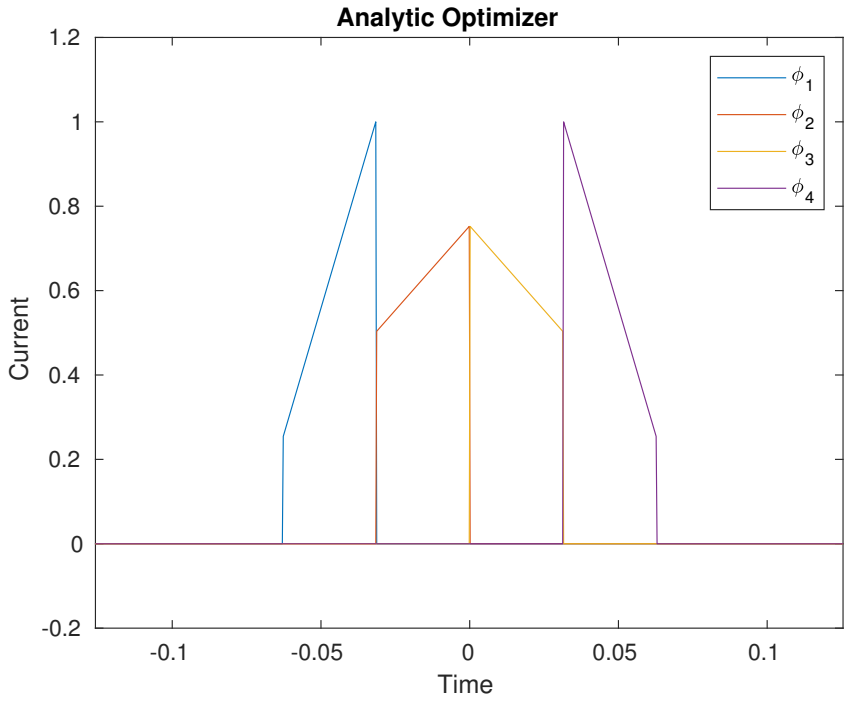

FIGURE 1.4: Partition of the maximizer ϕ

Moreover, $\hat{\psi}$ only vanishes on those points and it zeros are all simple. Therefore the ratio

$$\hat{\eta} := \frac{\hat{\phi}}{\hat{\psi}} \tag{1.31}$$

is well-defined. If $\hat{\eta}$ can be inverse Fourier transformed, we could write $\phi = \psi * \eta$, where η is the inverse Fourier transform of $\hat{\eta}$.

Proposition 1.9. *The optimal wavelet ϕ defined in (1.30) can be written as*

$$\phi = \eta * \psi \tag{1.32}$$

where ψ in the double sinc function defined in (1.14) and η is defined as

$$\begin{aligned}
\eta &= c_0 + c_1(\theta_\Delta + \theta_{-\Delta}) - c_2(\theta_\Delta - \theta_{-\Delta})D \qquad (1.33)\\
c_0 &= 2c\sin^2(\Delta)(2 + 2\cos(2\Delta))\\
c_1 &= 2c\sin^2(\Delta)\cos(\Delta)(1 + 2\cos^2(\Delta))\\
c_2 &= 2c\sin^2(\Delta)\sin(\Delta)\cos(2\Delta)
\end{aligned}$$

where c is the same as in (1.30), $\theta_\Delta(u)(t) := u(t-\Delta)$ is the shift operator and $Du(t) := u'(t)$ is the differential operator.

Proof. See appendix A.3 for explicit calculation. □

Though not totally unexpected, it is still a bit surprising that ϕ can be written as $\psi * \eta$, since there is no guarantee that $\hat{\eta}$ has inverse Fourier transform. Given this result, we conjecture that

Conjecture 1.10. *If ϕ is an optimizer of problem 1.23 without the constraint on $|\Omega_\phi|$, then $\phi = \psi * \eta$ almost surely for some η in the weak L^1 space.*

1.4.3 Numerical Convolution Optimizer

The efficiency of the optimal wavelet ϕ achieve is around 0.94, much greater than that of double sinc ψ, whose efficiency is around 0.75. The gain is due to the relaxation of the support size: the support size of ϕ is 4Δ while the support size of ψ is 2Δ, and extending the upper bound of $|\Omega_\phi|$ will further improve the efficiency. However, when the upper bound exceeds 4Δ the degree of freedom exceeds the number of constraints and it becomes difficult to solve problem 1.23 analytically. Thus we turn to numerical optimization through a convolution approach.

The goal of this section is to construct a generalization of 1.23 where the support bound can be $2l\Delta$ for any $l \in \mathbb{N}$ while the problem can still be solved numerically; this

can be achieved by reducing the constraints and approximating the objective function. Consider

$$\max_{\phi} E_2(\phi) \quad \text{subject to} \quad U_{\phi}^{\pm} = 1 \text{ and } |\Omega_{\phi}| \leq 2l\Delta. \tag{1.34}$$

for any $l \in \mathbb{N}$. The first step is motivated by conjecture 1.10. Let ψ be the double sinc function. If the conjecture holds, all solutions should take the form $\phi = \psi * \eta$ for some η in weak L^1, and the condition $U_{\psi}^{\pm} = 1$ is satisfied automatically. Thus we can re-write the problem as

$$\max_{\eta_l} E_2(\psi * \eta_l) \quad \text{subject to} \quad |\Omega_{\psi * \eta_l}| \leq 2l\Delta. \tag{1.35}$$

Along with the fact that $|\Omega_{\psi * \eta}| \sim |\Omega_{\psi}| + |\Omega_{\eta}|$, we can further reduce it into

$$\max_{\eta_l} E_2(\psi * \eta_l) \quad \text{subject to} \quad |\Omega_{\eta_l}| \leq 2(l-1)\Delta. \tag{1.36}$$

The second step is to approximate the objective function according to Holder inequality. Optimizers of problem 1.34 exist since E_2 is bounded above by 1 due to Holder inequality:

$$||\phi||_1 \leq ||\phi||_2 ||\chi_{\Omega_\phi}||_2 = ||\phi||_2 \sqrt{|\Omega_\phi|} \implies E_2(\phi) \leq 1. \tag{1.37}$$

Equation (1.37) implies that $E_2(\phi)$ is upper-bounded by 1 and equals to 1 if and only if ϕ is an indicator function almost surely. Although indicator functions do not satisfy our accuracy constraint (1.6), and hence not feasible, we would like to have our wavelet as close to an indicator function as possible. Therefore, we propose this alternative problem

$$\min_{\eta_l} ||\psi * \eta_l - \chi_{[-l\Delta, l\Delta]}||_2^2 \quad \text{subject to} \quad |\Omega_{\eta_l}| \leq 2(l-1)\Delta \tag{1.38}$$

which is simply a least square problem with a sparsity constraint that can be solved efficiently numerically, for any $l \in \mathbb{N}$. We call the solution of this problem the l-th "convolution optimizer" $\phi_l := \psi * \eta_l$.

In our numerical experiments, $E(\phi_l)$ increases monotonically with respect to l and

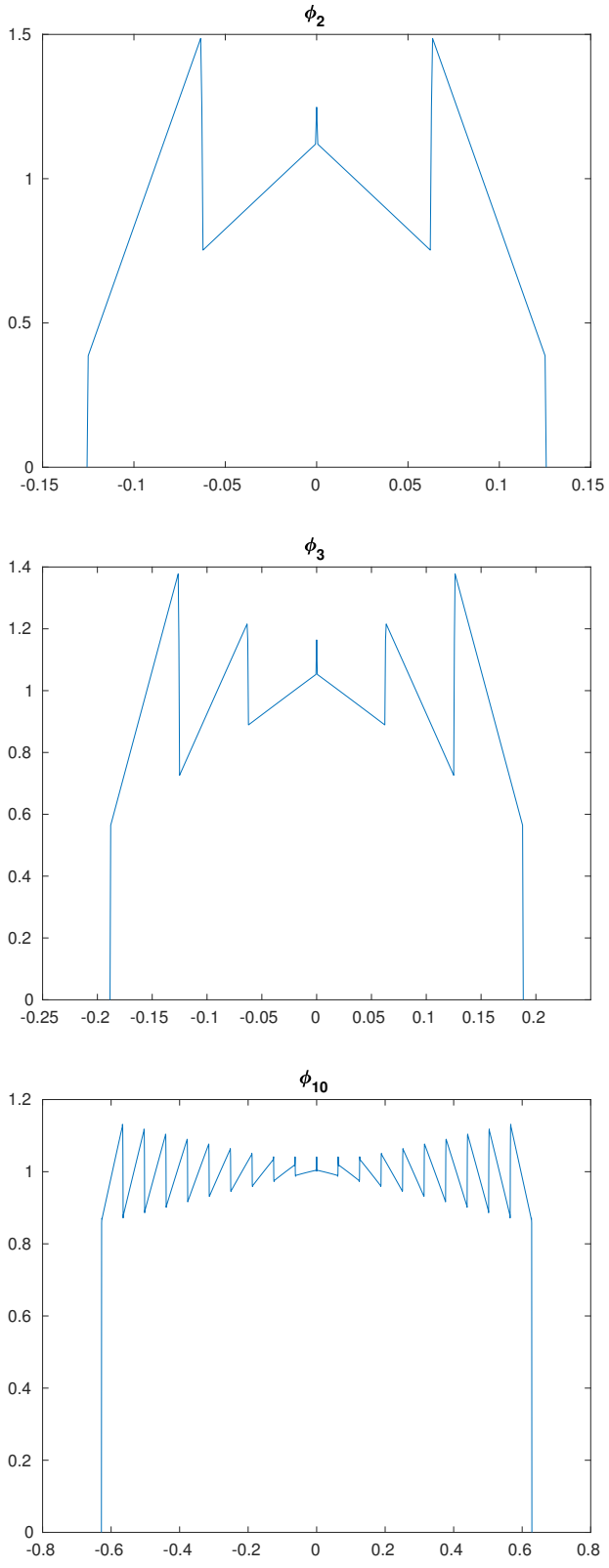

FIGURE 1.5: Structure of convolution optimizers $\phi_2, \phi_3, \phi_{10}$ (from left to right). Note that ϕ_2 is very similar to the analytic solution in figure 1.3.

l	1	2	3	4	5	6	7	8
$E(\phi_l)$	0.7538	0.9388	0.9729	0.9848	0.9903	0.9933	0.9951	0.9963

TABLE 1.1: Efficiency of convolution optimizers ϕ_l

approaches 1 asymptotically. According to table 1.1, the most significant increase occurs between $l = 1$ and $l = 2$, where we already have analytic expression (1.30). Figure 1.5 shows that when $l = 2$, we recover something close to what we derive in section 1.4.1. This suggests some consistency between the numerical approach (1.38) and the analytic approach (1.25). Moreover, the patterns in figure 1.5 provides some evidence of our conjecture 1.10.

As a side note, the middle spikes in figure 1.5 is a matter of numerical artifact that is bounded above and only occurs at a single point, which has negligible contribution to the efficiency.

1.5 Wavelet Comparison

We compare the efficiency of signal fragmentation using three different wavelets. Recall the definition of efficiency

$$E(\phi) = \frac{||\phi||_1^2}{|\Omega_\phi| \cdot ||\phi||_2^2}.$$

1.5.1 Truncated Gaussian

A frequently used wavelet is ϕ_{Gauss}, which is a truncated (for compact support) and shifted (for continuity) Gaussian, defined as

$$\phi_{\text{Gauss}}(t) =: \begin{cases} e^{-\frac{t^2}{2\sigma^2}} - e^{-\frac{\Delta^2}{2\sigma^2}} & \text{if } |t| \leq \Delta \\ 0 & \text{otherwise.} \end{cases} \tag{1.39}$$

Denote $\alpha = \frac{\Delta}{\sigma}$, then a direct calculation shows that

$$E(\phi_{\text{Gauss}}(\alpha)) = \frac{1}{2\alpha} \frac{\left(\sqrt{2\pi}\text{erf}\left(\frac{\alpha}{\sqrt{2}}\right) - 2e^{-\frac{\alpha^2}{2}}\alpha\right)^2}{\sqrt{\pi}\text{erf}(\alpha) - 2e^{-\frac{\alpha^2}{2}}\sqrt{2\pi}\text{erf}\left(\frac{\alpha}{\sqrt{2}}\right) + 2e^{-\alpha^2}\alpha} \tag{1.40}$$

in which $\text{erf}(x) := \frac{1}{\sqrt{\pi}}\int_{-x}^{x} e^{-\tau^2} d\tau$ is the error function. Note that (1.40) only depends on α. The maximal efficiency is around 0.83, which is achieved when α is small. Although ϕ_{Gauss} does not satisfy equation (1.7) and hence cannot perfectly reconstruct a sinusoidal signal, we nevertheless set 0.83 as a bar for our efficiency.

1.5.2 Double Sinc Function

The L^1 and L^2 norms and the efficiency E of the double sincc function are

$$||\psi||_1 = 2(1-\cos(\Delta)), \quad ||\psi||_2^2 = \Delta - \frac{1}{2}\sin(2\Delta)$$

$$E(\psi) = \frac{1}{2\Delta}\frac{4(1-\cos(x))^2}{\Delta - \frac{1}{2}\sin(2\Delta)} \approx \frac{4\frac{1}{4}\Delta^4}{2\Delta(\frac{1}{2}\frac{8}{6}\Delta^3)} = \frac{3}{4}. \tag{1.41}$$

Compared to ϕ_{Gauss}, ψ has higher accuracy for band-limited functions but lower efficiency.

1.5.3 Optimal Wavelet

Consider our optimal wavelet $\phi = \psi * \eta$. The L^1 and L^2 norms and the efficiency E are calculated (and approximated for large N) as

$$\begin{aligned}
||\psi*\eta||_1 &= (2c_0 + 4c_1)(1-\cos(\Delta)) \approx 2\pi \\
||\psi*\eta||_2^2 &= (\frac{1}{2}c_0^2 + c_1^2)(2\Delta - \sin(2\Delta)) + 4c_2^2\Delta \\
&\quad -2c_0c_1(\cos(\Delta)\Delta - \sin(\Delta)) - 4c_0c_2(1-\cos(\Delta)) \\
&\approx \frac{16\pi^2}{15}\Delta^{-1} \\
E(\psi*\eta) &= (4\Delta)^{-1}\frac{||\psi*\eta||_1^2}{||\psi*\eta||_2^2} \approx \frac{15}{16}
\end{aligned} \tag{1.42}$$

which has efficiency that is much higher than both ψ and ϕ_{Gauss}, and close to the optimal efficiency of 1.

1.6 Conclusion

In this chapter we quantify the projection error (spectral leakage) through the Poisson summation formula, establish its relation with the support size of wavelets, and optimize the efficiency under those constraints. We have analytically derive the double sinc function ψ, a wavelet that creates no spectral leakage while reconstructing sinusoidal wave and has the smallest support. Based on ψ, we further discover the analytic efficiency optimizer ϕ. In addition to no leakage, it also maximizes the efficiency under the constraint of its support size being less than 4Δ. Its efficiency is around 0.94, fairly close to the upper bound 1.

There are two possible directions for further work: one is to investigate the case where the order of zeros on lattices exceeds one, while the other is to prove the our conjecture 1.10. ψ is derived by taking $L = 1$ in theorem 1.3, but one may consider taking $L \in \mathbb{N}$ in general, and come up with a systematic way of constructing those wavelets with minimal support. The convolution conjecture, based on the observation that $\phi = \psi * \eta$, might require some specific regularity constraints since the efficiency needs to be well-defined.

Chapter 2

Superposed Multi-Reference Alignment Problem

In this chapter we propose a method that uses rational Fourier monomial and compressed sensing to recover signals from noisy observations. Let $x \in \mathbb{R}^N$ be the underlying signal and denote $x(n)$ to be the n-th entry of x, for $n \in [N] := \{0, \ldots, N-1\}$. Let G be a cyclic group of order N generated by z. We define the group action z^l acting on x as

$$(z^l \cdot x)(n) = x\left([n+l]_N\right) \tag{2.1}$$

where z^l can be viewed as the circular-shift operators. Here the notation $[n]_N$ is defined as $[n]_N := n \,(\text{mod } N)$.

Let g be a distribution over the group G. Our goal is to recover the underlying signal x from i.i.d. noisy samples of the form

$$y_i = g_{i,1} \cdot x + g_{i,2} \cdot x + \sigma\xi_i \tag{2.2}$$

where $g_{i,1}, g_{i,2} \sim g$, $\xi_i \sim \mathcal{N}(0, I_{N\times N})$ is the N-dimensional standard normal distribution, and all $g_{i,1}, g_{i,2}, \xi_i$ are independent. This is called the superposed multi-reference alignment problem (SMRA), an extension of multi-reference alignment problem (MRA), where

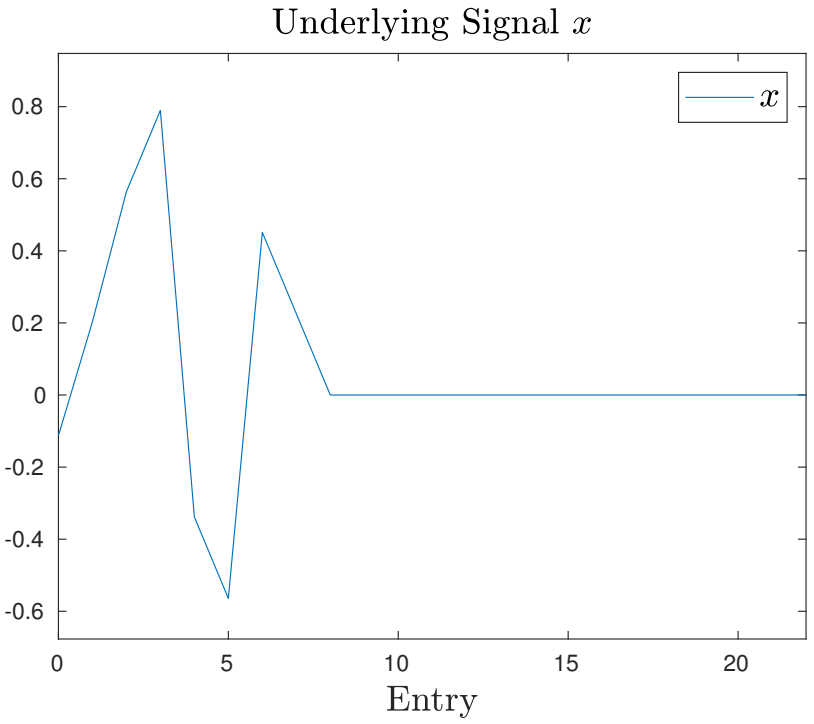

FIGURE 2.1: The underlying signal x

samples take the simpler form

$$y_i = g_i \cdot x + \sigma \xi_i. \tag{2.3}$$

Figure 2.2 illustrates how a sample y_i is generated from the ground truth shown in figure 2.1. In many applications, it suffices to recover an element in the orbit $G \cdot x$ because we care about the shape of the signal rather than its location/orientation. This objective is simpler than recovering x since we have enlarged our solution space.

We will provide motivation in section 2.1, introduce MRA techniques in section 2.2 and apply it to SMRA in section 2.3. We will then propose an algorithm based on compressed sensing in section 2.4, show some numerical results in section 2.5 and state our conclusion in section 2.6. Notation-wise, we define the discrete Fourier transform $\hat{x}$ or $\mathcal{F}(x)$ by

$$\hat{x}(k) = \frac{1}{N} \sum_{n=0}^{N-1} x(n) e^{-ink\Delta} \tag{2.4}$$

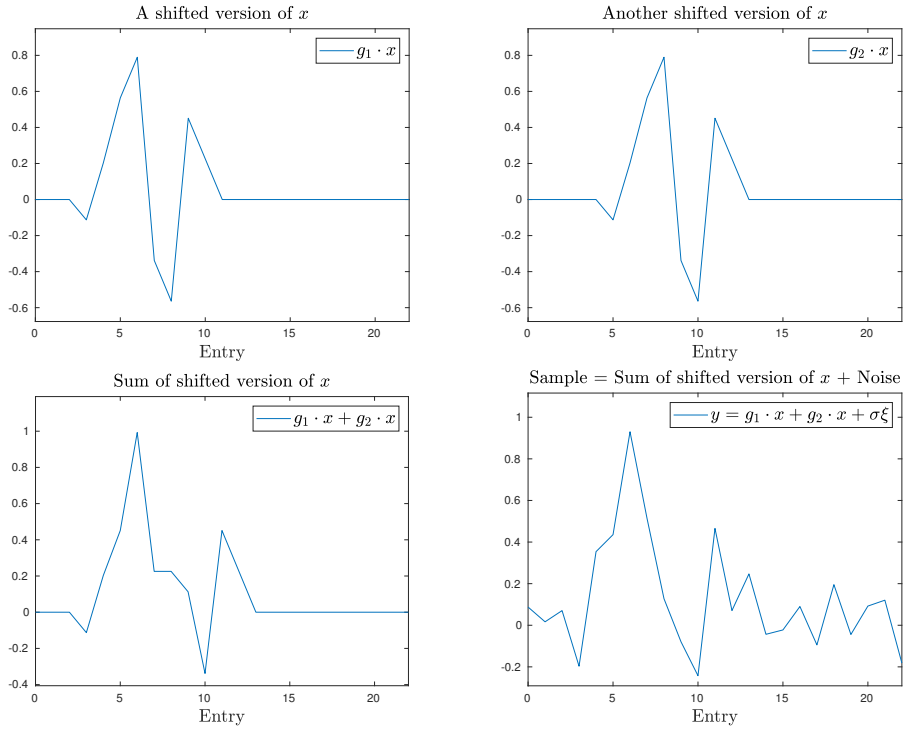

FIGURE 2.2: Illustration of samples y generated based on x

where $\Delta = \frac{2\pi}{N}$ and $k \in \mathbb{Z}$. With slight abuse of notation, in the case where g is a random shift operator, we denote

$$\mathcal{F}(g \cdot x) = e^{ig\Delta}\mathcal{F}(x) \tag{2.5}$$

where $e^{ig\Delta}$ is a random variable distributed over $e^{i[N]\Delta}$, since circular-shifting a signal is equivalent to multiplying its Fourier transform by a phase. Although the g has slightly different meaning in $g \cdot x$ and in $e^{ig\Delta}$, they are homomorphic and hence can represent each other.

2.1 Motivation: MRA and Cryo-EM

Signals are often corrupted by more than one type of noise, which increases the difficulty in signal recovery. For example, in one-dimensional multi-reference alignment problem (MRA), signals are randomly shifted and corrupted with Gaussian noise; in cryogenic

electron microscopy (Cryo-EM) [7], images of particles are randomly rotated and corrupted with Gaussian noise. We are interested in the cases where there are two types of noise: the regular Gaussian noise, denoted as ξ, and the noise that comes random group action, denoted as g, distributed over some known group G. In MRA, G is the cyclic group, corresponding to random circular shifts; in Cryo-EM, G is the special orthogonal group in dimension three, representing random three-dimensional rotations.

The goal is to recover x (or an element in$G \cdot x$) from samples y_i of the form

$$y_i = g_i \cdot x + \sigma \xi_i \tag{2.6}$$

where $g_i \sim g$ for some distribution g over G, ξ_i are Gaussian, and all g_i, ξ_i are independent.

This problem is nontrivial only when $\sigma > 0$ and g is not deterministic. It is clear that if $\sigma = 0$, one sample is sufficient. In the case where g is deterministic and equals to some element $\tilde{g} \in G$, we can recover the element $\tilde{g} \cdot x$ by averaging M samples since

$$\frac{1}{M}\sum_{i=1}^{M} y_i = \tilde{g} \cdot x + \frac{1}{M}\sum_{i=1}^{M} \sigma \xi_i \xrightarrow[M\to\infty]{} \tilde{g} \cdot x \in G \cdot x \tag{2.7}$$

by law of large numbers. Moreover, according to the central limit theorem, the variance of such estimation is $O(\sigma^2 M^{-1})$. However, in the case where $\sigma > 0$ and g is not deterministic, the same stradegy will not work because

$$\frac{1}{M}\sum_{i=1}^{M} y_i = \frac{1}{M}\sum_{i=1}^{M} (g_i \cdot x + \sigma \xi_i) \xrightarrow[M\to\infty]{} \mathbb{E}(g \cdot x) \notin G \cdot x. \tag{2.8}$$

illustrated by figure 2.5.

In section 2.1.1 and 2.1.2, we will review some classical techniques used for MRA, whose performance highly depends on the noise level σ. In section 2.2, we will introduce a recent mechanism that works for arbitrary noise level, and apply it to SMRA in section 2.3.

2.1.1 Pairwise Alignment

In the case where the Gaussian noise is small, i.e. $\sigma \ll 1$, it was shown that pairwise synchronization/alignment produces estimation that is positively correlated with some element in the orbit [8]. The idea is to estimate the relative "phase" $g_j g_i^{-1}$, and construct the aligned samples z_i (figure 2.3) from the original samples y_i (figure 2.4). For instance, fix $j \in \mathbb{N}$ and let

$$g_{ji} := \arg\min_{g \in G} ||g \cdot y_i - y_j|| \tag{2.9}$$

be the approximation of $g_j g_i^{-1}$ for all i. The aligned samples take the form

$$z_i := g_{ji} \cdot y_i \approx g_j \cdot x + \sigma \xi_i. \tag{2.10}$$

If the approximation is close to correct, then by similar calculation in (2.7), the element $g_j \cdot x$ can be approximated by the average

$$\frac{1}{M} \sum_{i=1}^{M} z_i \approx g_j \cdot x + \frac{1}{M} \sum_{i=1}^{M} \sigma \xi_i \xrightarrow[M \to \infty]{} g_j \cdot x \in G \cdot x \tag{2.11}$$

illustrated by figure 2.6.

However, the relative phase estimations are not exact and become less accurate as σ increases. When $\sigma > 1$, this method reaches its information theoretic limit [9], in the sense that the correlation between our estimation and any element in $G \cdot x$ is zero in average, that is, this method is no better than a random guess.

2.1.2 Frequency Marching

One alternative way is frequency marching [10], where instead of recovering the signal at once, we perform frequency estimation in a bounded domain, and extends its range at each step. It starts with estimating a set of low frequencies $\{\hat{x}(k)\}_{k=0,\dots,K}$ because they are less sensitive to noise. For instance, $\hat{x}(0)$ equals to the average of entries of x, which

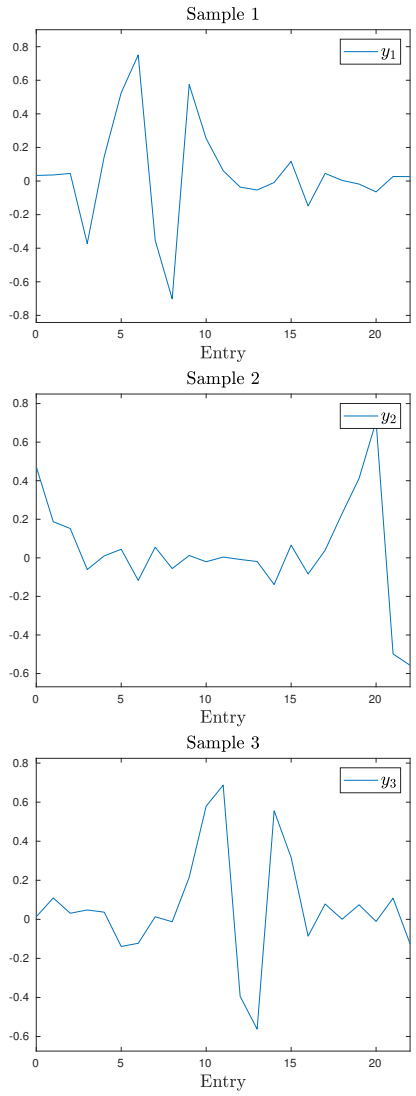

FIGURE 2.3: Original samples y_i

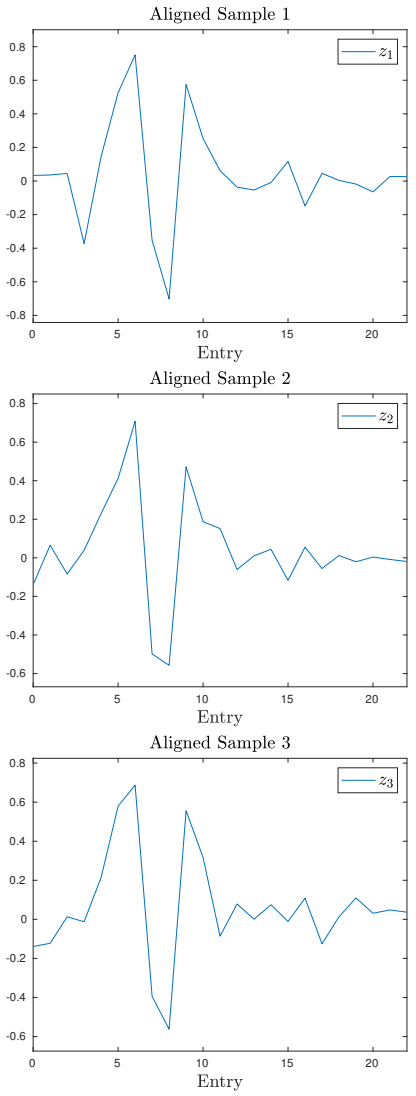

FIGURE 2.4: Aligned samples z_i

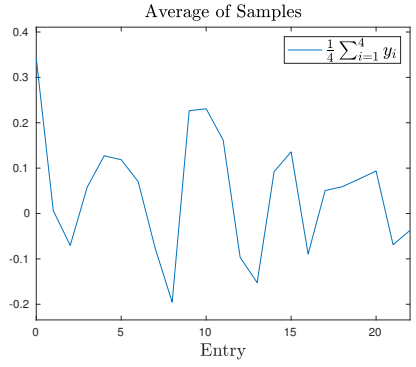

FIGURE 2.5: Averaging over original samples

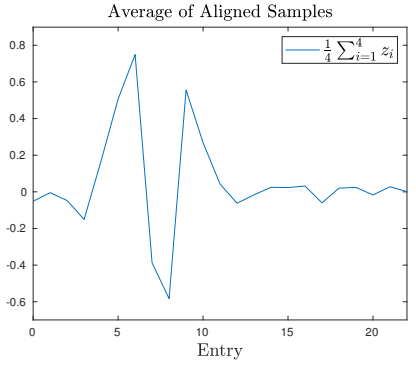

FIGURE 2.6: Averaging over aligned samples

is easy to estimate if the group action is Harr-distributed. We then use $\{\hat{x}(k)\}_{k=0,\dots,K}$ to estimate $\hat{x}(K+1)$. When a new frequency is introduced, the old ones will be modified accordingly so that the overall error is minimized. We carry this inductive procedure until we recover the full signal. Successful numerical results in Cryo-EM are shown in [10], but it is also pointed out that random initialization is subject to occasional failure, since this method requires solving multiple non-convex optimization problems and could be trapped by a local minimum.

2.2 Method of Invariant

Unlike the method of pairwise alignment and frequency marching, the method of invariant works for any $\sigma > 0$ and does not require an initial guess. The downside is that it requires more samples than others. For instance in MRA, the number of samples required for nontrivial estimation is $O(\sigma^2)$ for most algorithm, but the method of invariant requires $O(\sigma^6)$ samples because it involves estimation of the third moment of the signal [11]. Such scaling, though seem undesirable, is in fact optimal in the sense that when $\sigma \gg 1$, it is necessary and sufficient to use $O(\sigma^6)$ samples [9], [12].

The method of invariant consists of two part: finding invariant under the group action and recovering the signals from those invariant. We will examine how those techniques work in MRA, and use them to solve the superposed MRA problem, abbreviated as SMRA. To achieve such goal, we will first define the invariant in the general sense.

Definition 2.1. (Invariant)
Let y be a random variable related to an underlying signal $x \in \mathbb{R}^N$. We say a functional $f : \mathbb{R}^N \times \mathbb{R}^{N_1} \to \mathbb{R}$ for some $N_1 \in \mathbb{N}$ is an invariant of x with respect to y if

$$\mathbb{E}[f(y,k)] = c_1 f(x,k) + c_2 \tag{2.12}$$

for some deterministic c_1 and c_2 and the variance $\mathrm{Var}[f(y,k)]$ is finite for all $k \in [N_1]$. For simplicity we will denote $f(y,k)$ with $f_y(k)$.

In other words, invariant determine the types of partial information of x that we can deduce from y. For example in MRA, the first and second moment of a vector, $f(x,k) := \sum_{n\in[N]} x^k(n)$, $k = 1, 2$, are invariant of x with respect to y because

$$\mathbb{E}[f(y,1)] = \sum_{n\in[N]} x(n) + \sigma\mathbb{E}[\xi(n)] = f(x,1)$$
$$\mathbb{E}[f(y,2)] = \sum_{n\in[N]} x^2(n) + 2\sigma x(n)\mathbb{E}[\xi(n)] + \sigma^2\mathbb{E}[\xi^2(n)] = f(x,2) + \sigma^2.$$

Among all invariant, we would like to find the ones that help us recover the original signal. In MRA, they are the spectrum and bi-spectrum: lemma 2.2 states that spectrum and bi-spectrum are invariant, while theorem 2.3 claims that they are sufficient for recovery of x. Intuitively, since the spectrum contains the information about the magnitudes and the bi-spectrum provides the information about the phases, we can fully recover $\hat{x}$ and hence x, assuming the $\hat{x}$ does not vanish at any point. We will apply similar techniques to SMRA in the later sections.

Lemma 2.2. *In MRA, the spectrum S and bi-spectrum B defined as*

$$S_x(k) := \hat{x}(k)\hat{x}(-k) \tag{2.13}$$
$$B_x(k_1,k_2) := \hat{x}(k_1)\hat{x}(k_2)\hat{x}(-k_1-k_2) \tag{2.14}$$

are invariant of x with respect to y.

Proof. See B.1. □

Theorem 2.3. *A signal $x \in \mathbb{R}^N$ with $\hat{x}(k) \neq 0$ for all $k \in [N]$ can be fully recovered from its spectrum and bi-spectrum up to a global shift.*

Proof. By assumption, $S_x(k)$ and $B_x(k_1,k_2)$ are also nonzero for all $k, k_1, k_2 \in \mathbb{Z}$. We will first compute $\hat{x}(0)$ and $\hat{x}(1)$ through S and B and finish the proof by induction. The

calculation of $\hat{x}(0)$ is relatively straightforward:

$$\hat{x}(0) = \frac{B_x(0,0)}{S_x(0)}. \tag{2.15}$$

By periodicity, $\prod_{k=0}^{N-1} \hat{x}(k) = \prod_{k=0}^{N-1} \hat{x}(k+n)$ for any $n \in \mathbb{Z}$ and hence

$$\begin{aligned}
\prod_{k=0}^{N-1} B_x(1,k) &= \prod_{k=0}^{N-1} \hat{x}(1)\hat{x}(k)\hat{x}(-1-k) \\
&= [\hat{x}(1)]^N \prod_{k=0}^{N-1} \hat{x}(k) \prod_{k=0}^{N-1} \hat{x}(-1-k) \\
&= [\hat{x}(1)]^N \prod_{k=0}^{N-1} \hat{x}(k) \prod_{k=0}^{N-1} \hat{x}(-k) \\
&= [\hat{x}(1)]^N \prod_{k=0}^{N-1} S_x(k)
\end{aligned}$$

which enables us to compute $\hat{x}(1)$ by

$$\hat{x}(1) = \left[\prod_{k=0}^{N-1} \frac{B_x(1,k)}{S_x(k)} \right]^{\frac{1}{N}}. \tag{2.16}$$

Note that there are N possible values $\hat{x}(1)$ can take as we take the N-th root, each of them corresponding to a different global shift of x. For our purpose this ambiguity will not affect the outcome, so we will pick one at random. Given $\hat{x}(0)$ and $\hat{x}(1)$, we may recover other frequencies by the following inductive steps

$$\hat{x}(-k) = \frac{S_x(k)}{\hat{x}(k)} \tag{2.17}$$

$$\hat{x}(k+1) = \frac{B_x(-1,-k)}{\hat{x}(-1)\hat{x}(-k)} \tag{2.18}$$

until we fully recover $\hat{x}$ and hence x up to a global shift. □

The significance of theorem 2.3 is that it not only specifies the method for signal recovery, but also provides the number of samples required for nontrivial estimation,

which is proportional to the variance of our invariant. Since the variance of spectrum and bi-spectrum are $O(\sigma^4)$ and $O(\sigma^6)$, the number of samples required is $O(\sigma^6)$ [13].

2.3 SMRA

Recall that in SMRA our objective is to recover the signal x from samples of the form

$$y = g_1 \cdot x + g_2 \cdot x + \sigma\xi \tag{2.19}$$

where g_1, g_2 are random shifts and ξ are Gaussian. We will examine the well-posedness of this problem, find its invariant, and propose an algorithm for recovery.

2.3.1 Well-posedness

Even without the presence of Gaussian noise, SMRA might not be well-posed due to cancellation between $g_1 \cdot x$ and $g_2 \cdot x$, which is easier to analyzed in Fourier space. In particular if $y = g_1 \cdot x + g_2 \cdot x$, then the k-th frequency

$$\hat{y}(k) = \left(e^{ig_1 k\Delta} + e^{ig_2 k\Delta}\right)\hat{x}(k) = c(k)\hat{x}(k)$$

contains no information of x if $c(k) = 0$. According to lemma 2.4, which also works for the general sum $y = \sum_{m=1}^{M} g_m \cdot x$, we can avoid such issue by requiring N to be a prime.

Lemma 2.4. *If N is a prime, $N > M \in \mathbb{N}$, and $y = \sum_{m=1}^{M} g_m \cdot x$, then $\hat{y}(k) = 0$ if and only if $\hat{x}(k) = 0$.*

Proof. Let $\Delta = \frac{2\pi}{N}$. By definition

$$\hat{y}(k) = \left(\sum_{m=1}^{M} e^{ig_m k\Delta}\right)\hat{x}(k) = c(k)\hat{x}(k)$$

and hence it suffices to show that $c(k) \neq 0$ for all $k \in [N]$. We will prove this by contradiction. Suppose $c(k) = 0$ for some $k \in [N]$. Since $e^{igk\Delta} = e^{i[gk]_N\Delta}$, let $a_n = |\{m : [g_m k]_N = n\}|$ be the number of terms in $c(k)$ that equals to $e^{in\Delta}$. Then $c(k) = 0$ implies that

$$0 = \sum_{m=1}^{M} e^{ig_m k\Delta} = \sum_{m=1}^{M} e^{i[g_m k]_N\Delta} = \sum_{n=0}^{N-1} a_n e^{in\Delta} \tag{2.20}$$

which means that the polynomial

$$p(z) := \sum_{n=0}^{N-1} a_n z^n \tag{2.21}$$

has a root at $e^{i\Delta}$. Note that $a_n \in \mathbb{Z}$ and some coefficients a_n in p must be zero because $N > M$ by assumption. Since N is a prime, the N-th cyclotomic monomial

$$\Phi_N := \sum_{n=0}^{N-1} z^n \tag{2.22}$$

is a minimal monic polynomial over the field of the rational numbers. According to the irreducibility of Φ_N (see B.2), the greatest common divisor of Φ_N and p is the constant function 1. By Bézout's theorem, there exists polynomials a and b with rational coefficients such that

$$(ap + b\Phi_N)(z) = \gcd(p, \Phi_N)(z) = 1. \tag{2.23}$$

Plug in $z = e^{i\Delta}$, we get $0 = 1$, hence the contradiction. □

Moreover, we need to separate the case where $g_1 = g_2$ and $g_1 \neq g_2$ due to the ambiguity illustrated in figure 2.7, which is justified by proposition 2.5 as we explore the algebraic structure of SMRA.

Proposition 2.5. *If N is a prime, $y = g_1 \cdot x + g_2 \cdot x$, then the total Fourier product*

$$T_x := \prod_{k=0}^{N-1} \hat{y}(k) \tag{2.24}$$

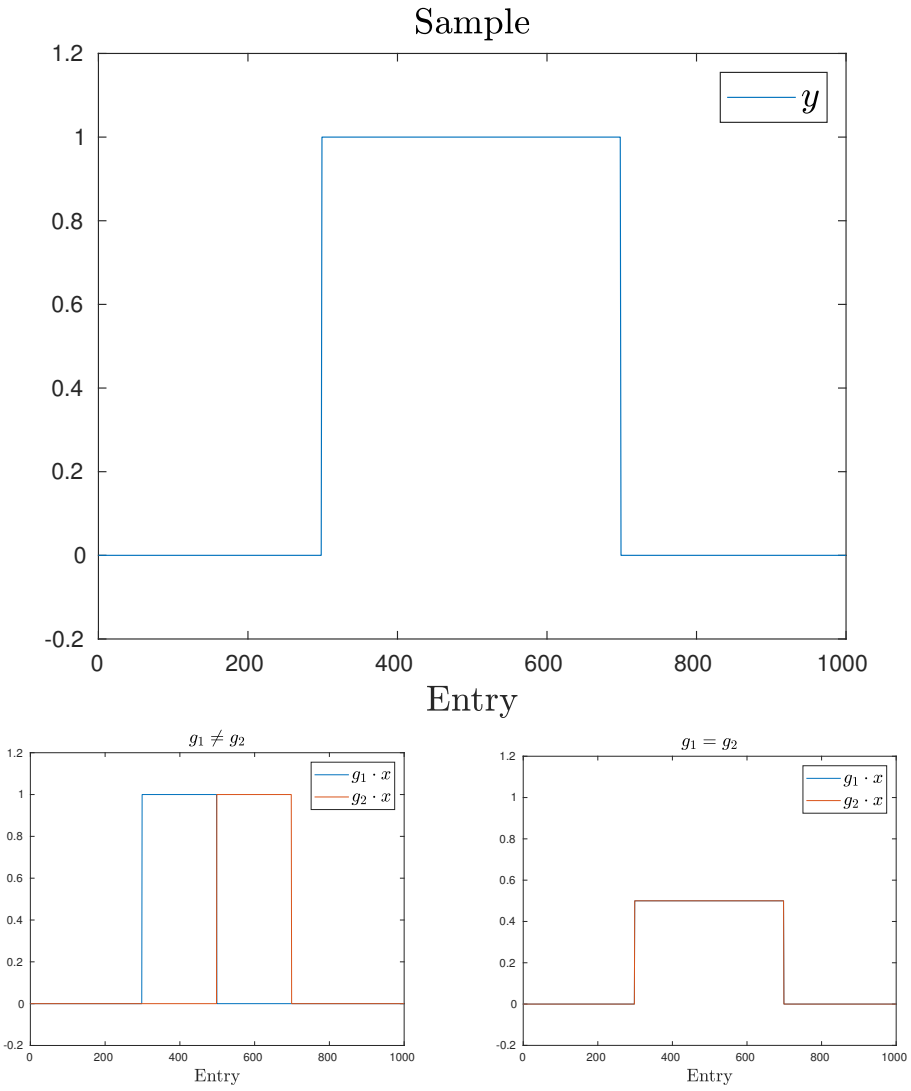

FIGURE 2.7: A superposed signal (top) that allows different decomposition, depending on whether the shift is the same (right) or not (left)

satisfies the following equation

$$T_y = cT_x \tag{2.25}$$

where $c = 2$ *if* $g_1 \neq g_2$ *and* $p = 2^N$ *if* $g_1 = g_2$.

As an interesting side step, this proposition leads to a proof of Fermat's Little theorem in the case of $n = 2$. See appendix B.4.

Proof. Suppose $g_1 \neq g_2$ and let $\Phi_N(z) := z^N + 1 = \prod_{k=0}^{N-1}(z + e^{ik\Delta})$ be the N-th cyclotomic polynomial. By the algebraic structure of primary cyclic group,

$$\begin{aligned} c &= \prod_{k=0}^{N-1}(e^{ig_1k\Delta} + e^{ig_2k\Delta}) = \prod_{k=0}^{N-1}(1 + e^{i(g_2-g_1)k\Delta}) \\ &= \prod_{k=0}^{N-1}(1 + e^{ik\Delta}) \quad \text{(since } N \text{ is a prime and } g_1 \neq g_2\text{)} \\ &= \Phi_N(1) = 2. \end{aligned}$$

The case of $g_1 = g_2$ is trivial. □

For well-posedness, we will assume N is a prime and $g_1 \neq g_2$ in the following sections. Although g_1 and g_2 are no longer independent under the restriction $g_1 \neq g_2$, they are almost independent in the sense that the new joint distribution $\mathbb{P}(g_1 = a, g_2 = b)$ is defined as

$$\mathbb{P}(g_1 = a, g_2 = b) = \begin{cases} c\mathbb{P}(g_1 = a)\mathbb{P}(g_2 = b) & \text{if } a \neq b \\ 0 & \text{otherwise} \end{cases} \tag{2.26}$$

for some normalizing constant c.

2.3.2 Invariant: Rational Fourier Monomial

In this section we define and analyze some invariant for SMRA. To better illustrate our intuition, we will set $\sigma = 0$ temporally. Observe that S and B are not invariant in SMRA because

$$\begin{aligned} S_y(k) &= (e^{ig_1k\Delta} + e^{ig_2k\Delta})\hat{x}(k)(e^{-ig_1k\Delta} + e^{-ig_2k\Delta})\hat{x}(-k) \\ &= (2 + 2\cos((g_1 - g_2)k\Delta))S_x(k) \end{aligned}$$

depend on $g_1 - g_2$. However, the rational Fourier monomial are invariant in SMRA when $\sigma = 0$.

Definition 2.6. (Rational Fourier Monomial)

Given $x \in \mathbb{R}^N$ such that N is a prime $\hat{x}(k) \neq 0$ for all $k \in [N]$, we define its rational Fourier monomials of degree $L \geq 3$ to be monomials of the form

$$P_x^L(k_1, \ldots, k_L) := \prod_{l=1}^{L} \frac{\hat{x}(k_l)}{\hat{x}(-k_l)} \tag{2.27}$$

such that $\sum_{l=1}^{L} k_l = 0$.

Proposition 2.7. *If N is prime and $\hat{x}(k) \neq 0$ for all $k \in [N]$, then in SMRA when $\sigma = 0$,*

$$P_y^L(k_1, \ldots, k_L) = P_x^L(k_1, \ldots, k_L) \tag{2.28}$$

Proof. By lemma 2.4 P_y^N is well-defined. Straightforward calculations show that

$$\begin{aligned}
R_y(k) := \frac{\hat{y}(k)}{\hat{y}(-k)} &= \frac{(e^{ig_1k\Delta} + e^{ig_2k\Delta})\hat{x}(k)}{(e^{-ig_1k\Delta} + e^{-ig_2k\Delta})\hat{x}(-k)} \\
&= e^{i(g_1+g_2)k\Delta} \frac{\hat{x}(k)}{\hat{x}(-k)} \\
\implies P_y^L(k_1, \ldots, k_L) &= \prod_{l=1}^{L} e^{i(g_1+g_2)k_l\Delta} \frac{\hat{x}(k_l)}{\hat{x}(-k_l)} \\
&= \left(\prod_{l=1}^{L} e^{i(g_1+g_2)k_l\Delta} \right) \prod_{l=1}^{L} \frac{\hat{x}(k_l)}{\hat{x}(-k_l)} \\
&= e^{i(g_1+g_2)\Delta \sum_{l=1}^{L} k_l} P_x^L(k_1, \ldots, k_L) \\
&= P_x^L(k_1, \ldots, k_L)
\end{aligned}$$

since $\sum_{l=1}^{L} k_l = 0$. □

Proposition 2.5 and 2.7 show that the total Fourier product T and rational Fourier monomials P^L are invariant without Gaussian noise. Unfortunately, rational Fourier monomials are not invariant when $\sigma > 0$ due to the fact that the expectation of a ratio

does not equal to the ratio of expectation. Here is a simple example: for $y = x + \xi$,

$$\mathbb{E}\left[\frac{\hat{y}(k)}{\hat{y}(-k)}\right] = \mathbb{E}\left[\frac{\hat{x}(k) + \hat{\xi}(k)}{\hat{x}(-k) + \hat{\xi}(-k)}\right] \neq \frac{\hat{x}(k)}{\hat{x}(-k)} = \frac{\mathbb{E}\left[\hat{y}(k)\right]}{\mathbb{E}\left[\hat{y}(-k)\right]}. \tag{2.29}$$

Though exact recovery may be impossible, under some assumptions we can write

$$\mathbb{E}\left[P_y^L(k_1, \ldots, k_L)\right] \approx c_1 P_x^L(k_1, \ldots, k_L) + c_2 \tag{2.30}$$

for some deterministic c_1, c_2. We will discuss those assumptions in section 2.3.3.

2.3.3 Ratio Estimator

Suppose $\sigma > 0$ and $y = g_1 \cdot x + g_2 \cdot x + \sigma\xi$. We would like to estimate our rational Fourier monomials from y. In general given two random variables X, Y with mean μ_X, μ_Y and variance σ_X, σ_Y, the ratio estimator with second order Taylor expansion (see appendix B.3 for details) has expectation

$$\mathbb{E}\left[\frac{X}{Y}\right] \approx (1 + \mu_Y^{-2}\sigma_Y^2)\mu_X\mu_Y^{-1} - \mu_Y^{-2}\mathrm{Cov}[X, Y] \tag{2.31}$$

and variance

$$\mathrm{Var}\left[\frac{X}{Y}\right] \approx \mu_Y^{-2}\sigma_X^2 + \mu_X^2\mu_Y^{-4}\sigma_Y^2 - 2\mu_X\mu_Y^{-3}\mathrm{Cov}[X, Y]. \tag{2.32}$$

which implies that

$$\begin{aligned}
\mathbb{E}\left[P_y^L(k_1, \ldots, k_L)\right] \approx & \left[1 + \prod_{l=1}^{L} \frac{1}{\hat{x}^2(-k_l)}\sigma^{2L}\right] P_x^L(k_1, \ldots, k_L) \\
& - \prod_{l=1}^{L} \frac{1}{\hat{x}^2(-k_l)}\mathrm{Cov}\left[\left(\prod_{l=1}^{L} \hat{y}^2(k_l)\right)\left(\prod_{l=1}^{L} \hat{y}^2(-k_l)\right)\right] \qquad (2.33) \\
= & \, c_1 P_x^L(k_1, \ldots, k_L) + c_2. \qquad (2.34)
\end{aligned}$$

Since $c_1 = c_1(x, k_1, \ldots, k_L)$ and $c_2 = c_2(x, k_1, \ldots, k_L)$ are not constant in $\mathbb{R}$, it can be quite complicated to estimate $P_x^L(k_1, \ldots, k_L)$ with $P_y^L(k_1, \ldots, k_L)$. We believe that there are statistical tools that can help us with this, and for now we will assume that we can estimate $P_x^L(k_1, \ldots, k_L)$.

2.4 Algorithm

In this section we introduce an algorithm that approximates the underlying signal with rational Fourier monomials and compressed sensing.

2.4.1 Phase and Spectrum Product Estimation

Analogous to the procedure in (2.16), (2.17) and (2.18), we can deduce the phase of $\hat{x}$ up to a negative sign by the following inductive steps: by definition $R_x(0) = 1$. Due to periodicity,

$$R_x(1) = \left[\prod_{k=0}^{N-1} P_x^3(1, k, -1-k)\right]^{\frac{1}{N}} = \left[\prod_{k=0}^{N-1} P_y^3(1, k, -1-k)\right]^{\frac{1}{N}} \tag{2.35}$$

up to a factor of $e^{in\Delta}$, $n \in [N]$, that corresponds to a global shift of x. Once $R_x(1)$ is determined, we can calculate the rest by the following induction

$$R_x(-k) = R_x(k)^{-1} \tag{2.36}$$

$$R_x(k+1) = P_y^3(-1, -k, k+1) R_x(1) R_x(k). \tag{2.37}$$

Since rational Fourier monomials only provide information of the phase, third-degree monomials seem to be enough. To gain control over the spectrum, we will use the fact that the total Fourier product is also an invariant.

In summary, we manage to estimate the following properties of x:

1. phases of $\hat{x}$ up to a negative sign, denoted as $e^{i\theta_k}$, and

2. total product of $\hat{x}$, denoted as T_x.

The estimation derived above is insufficient for recovering x exactly, but can used for approximation if we impose extra constraints such as sparsity. In section 2.4.2 we will introduce an algorithm from compressed sensing that helps us achieve such goal.

2.4.2 Compressed Sensing

Compressed sensing through L^1 regularization are widely used in image processing and computer vision, when the system is under-determined and the signal is sparse. It was proposed in [14] and further explored in [15]. It is suitable for our problem because the signals (images of molecules) are sparse over the domain (a plate). To apply the technique in compressed sensing, we formulate our signal recovery problem into the following optimization problem:

$$\min_{a_k \in \mathbb{R}} |\Omega_x| \quad \text{such that } \hat{x}(k) = a_k e^{i\theta_k}, \prod_{k=0}^{N-1} a_k = T_x \tag{2.38}$$

where $e^{i\theta_k}$ and T_x are given. We can write this in a more compact form by defining the N-by-N matrix M with $M_{mk} = \exp(imk\Delta + i\theta_k)$ and the vector $a \in \mathbb{R}^N$. Then we have

$$\min_{a \in \mathbb{R}^N} |\Omega_{Ma}| \quad \text{such that } \text{prod}(a) = T_x. \tag{2.39}$$

This falls into the category of the split Bregman method [15]. In the second step of algorithm 1, the minimization problem is solved through gradient descend under Wolfe condition. To avoid local minimum, a must be initialized in the right quadrant.

2.5 Numerical Results

In this section we focuses on the relation between the accuracy of our approximation and the sparsity of the signal, defined as the percentage of nonzero entries. We assume

Algorithm 1 Generalized Split Bregman

Input $M \in \mathbb{R}^{N\times N}, T_x \in \mathbb{R}$
Output $a \in \mathbb{R}^N$
Random Initialization $a, b, d \in \mathbb{R}^N$

while $||a^k - a^{k-1}||_2 >$tol **do**
  $a^{k+1} \leftarrow \min_a (\text{prod}(a) - T_x)^2 + \frac{\lambda}{2}||d^k - Ma - b^k||_2^2$
  $d^{k+1} \leftarrow \min_d |d| + \frac{\lambda}{2}||d^k - Ma^{k+1} - b^k||_2^2$
  $b^{k+1} \leftarrow b^k + (Ma^{k+1} - d^{k+1})$
end while

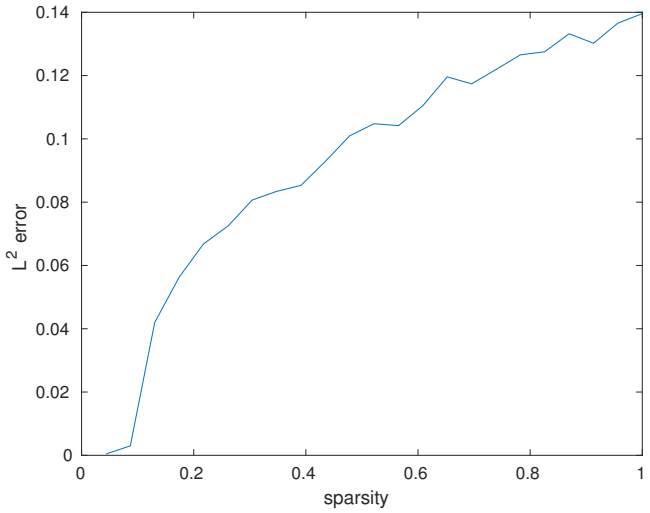

FIGURE 2.8: The relation between the l^2 error in our estimation and sparsity of the signal, normalized compared to random guesses

infinite samples so that all invariant can be accurately calculated, and hence the only error is due to our algorithm. In our numerical experiment $N = 23$. For each sparsity level $s \in ([N] + 1)/N$, we generate 400 random signals x_i, each with two random shifts $g_{i,1}, g_{i,2}$. We then apply algorithm 2 to each observation $y_i = g_{i,1} \cdot x + g_{i,2} \cdot x$ and compute the l^2 difference between the estimation $\tilde{x_i}$ and the signal x_i. The average and normalized (compare to random guess with $z \sim \mathcal{N}(0, I_{N\times N})$) error are shown in figure 2.8.

As an example, the result of applying algorithm 2 to samples in figure 2.2 is shown in figure 2.9. In general, the l^2 error increases as the sparsity coefficient increases (less nonzero entries), as illustrated by figure 2.8, but they are fairly accurate since the error is far away from 1, the mean error for random guess.

Algorithm 2 Signal recovery through rational Fourier monomial

Input $y \in \mathbb{R}^N$
Output $x \in \mathbb{R}^N$

procedure PHASE ESTIMATION
$R(0) \leftarrow 1$
$R(1) \leftarrow \left[\prod_{k=0}^{N-1} P_y^3(1, k, -1-k)\right]^{\frac{1}{N}}$,
for $k = 1, \ldots, N-1$ **do**
$R(-k) \leftarrow R(k)^{-1}$
$R(k+1) \leftarrow P_y^3(-1, -k, k+1)R(1)R(k)$
end for
end procedure

procedure TOTAL PRODUCT ESTIMATION
$T \leftarrow \frac{1}{2}$
for $k = 1, \ldots, N$ **do**
$T \leftarrow T \cdot \hat{y}(k)$
end for
end procedure

procedure COMPRESSED SENSING
Random initialization $a, b, d \in \mathbb{R}^N$
for $m = 1, \ldots, N$ and $k = 1, \ldots, N$ **do**
$M(m, k) = e^{imk} \cdot R(k)$
end for
while $||a^k - a^{k-1}||_2 >$tol **do**
$a^{k+1} \leftarrow \min_a (\text{prod}(a) - T)^2 + \frac{\lambda}{2}||d^k - Ma - b^k||_2^2$
$d^{k+1} \leftarrow \min_d |d| + \frac{\lambda}{2}||d^k - Ma^{k+1} - b^k||_2^2$
$b^{k+1} \leftarrow b^k + (Ma^{k+1} - d^{k+1})$
end while
end procedure

Output $x = Ma$

2.6 Conclusion

In this chapter we explore the algebraic structure of the sum of two noisy copies of a signal under random cyclic shifts. We have shown that in the discrete setting, when the order of the cyclic group is prime, phases of the signal's Fourier transformation can be determined up to a negative sign, and the total product of the noisy spectrum is proportional to that of the original signal. With further assumption on the support of the signal, we formulate an

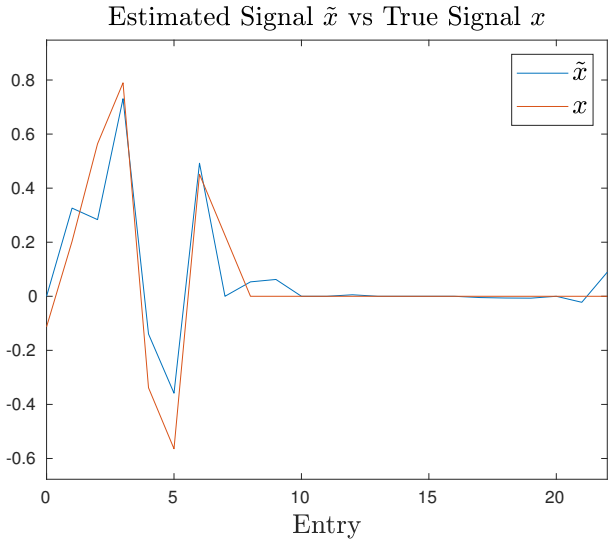

FIGURE 2.9: An example of estimation that deviates from the underlying signal but preserves the signal's general shape

optimization problem that can be solved through split Bergman algorithm. The average performance is quite impressive, assuming that we can estimate the third-degree rational Fourier monomial accurately. In the future we would like to add more analysis on this method and extend it to solving sum of M noisy copies of a signal.

Chapter 3

Unsteady Current Beyond Child-Langmuir Limit

In this chapter we will use Euler-Poisson equations to analyze the flow of electrons in a diode. In fluid dynamics, Euler equations are a set of quasilinear hyperbolic equations governing the dynamics of flows. Intuitively, Euler Equations can be understood as the conservation of density in time, i.e.

$$\partial_t \rho + u \cdot \nabla_x \rho + F \cdot \nabla_u \rho = 0 \tag{3.1}$$

where $\rho = \rho(u, x, t)$ is an distribution representing the density, u is the velocity, t is the time variable, x is the space variable, and F is a force term. In electrostatics, Poisson's equation states that the force F is related to the potential ϕ, which is related to the density ρ in the following way

$$F = \nabla_x \phi \tag{3.2}$$

$$\nabla_x^2 \phi = \rho. \tag{3.3}$$

With some manipulations of (3.1), (3.2), and (3.3), we arrive at the Euler-Poisson (EP) equations that governs the current of an electron beam in a diode:

$$\partial_t \rho + \partial_x(\rho u) = 0 \tag{3.4}$$

$$\partial_t u + u\partial_x u - \partial_x \phi = 0 \tag{3.5}$$

$$\partial_{xx}\phi - \rho = 0 \tag{3.6}$$

over the domain $\{(x,t) : (x,t) \in [0,L] \times [0,T]\}$, with boundary conditions

$$\rho(0,t) = \rho_0, \quad u(0,t) = u_0, \quad \phi(L,t) = \phi_L, \quad \phi(0,t) = 0 \tag{3.7}$$

and periodic boundary conditions

$$\rho(x,0) = \rho(x,T), \quad u(x,0) = u(x,T), \quad \phi(x,0) = \phi(x,T) \tag{3.8}$$

where ρ, u, ϕ represent velocity, density, and potential respectively. We use the notation ∂_x instead of ∇_x because we are working with one-dimensional-space, corresponding to the direction that perpendicular to the cathodes and anodes.

Physicists were interested in the space-charge limited current, i.e. the maximal current $J := \rho u$ given the boundary conditions in velocity u_0, potential ϕ_L, and the length L of the device. Such limit was derived for a simpler model by Child [16], where all variables are assumed to be time-independent and the Euler-Poisson equations are reduced to the time-independent Euler-Poisson equations:

$$\partial_x(\rho u) = 0 \tag{3.9}$$

$$u\partial_x u - \partial_x \phi = 0 \tag{3.10}$$

$$\partial_{xx}\phi - \rho = 0. \tag{3.11}$$

Note that equation (3.9) implies that the current

$$J := \rho u$$

is a constant. Child [16] and Langmuir [17] derived an upper bound for the current in the case where $u_0 = 0$, which is consistent with their experiment results. Later Jaffe [18] extended it to the case for general $u_0 \geq 0$. This limit, which we will call it the CL limit, is derived from the method of characteristics and defined as

$$J_{\mathrm{CL}}(u_0, \phi_L, L) = \frac{2}{9L^2}\left(u_0 + \sqrt{u_0^2 + 2\phi_L}\right)^3 \tag{3.12}$$

where full derivation will be shown in section 3.1.

However, such limit no longer holds in the original Euler-Poisson equations, where the solutions are time dependent. In particular, there are numerical evidence [19] [20] showing that there exist a periodic boundary condition (ρ_0, u_0, ϕ_L, L) corresponding to a solution whose average flux exceeds the CL limit, i.e.

$$\bar{J}_0(\rho_0, u_0) > \bar{J}_{\mathrm{CL}}(u_0, \phi_L, L) \tag{3.13}$$

where

$$\bar{J}_0(\rho_0, u_0) := \frac{1}{T}\int_0^T \rho_0 u_0 dt \tag{3.14}$$

$$\bar{J}_{\mathrm{CL}}(u_0, \phi_L, L) := \frac{1}{T}\int_0^T J_{\mathrm{CL}}(u_0, \phi_L, L) dt. \tag{3.15}$$

We call $\bar{J}_0$ the average flux rather than current because its definition involves integration. Our goal is to construct a systematic way in finding the maximizer of the flux difference

$$\bar{J}_{\mathrm{diff}} := \bar{J}_0 - \bar{J}_{\mathrm{CL}} \tag{3.16}$$

through variation principle.

In the following sections, we will first derive the CL limit in section 3.1, define objective Lagrangian and calculate its derivatives in section 3.2.1, and describe the numerical algorithm in section 3.3. In section 3.4 we will show some numerical results, including the flux difference and their characteristics.

3.1 CL Limit

In this section we derive the CL limit (3.12) with the method of characteristic and analyze its limitation.

3.1.1 Derivation

Suppose ρ, u, ϕ are time independent and denote $\psi(x) := \phi'(x)$ to be the electric field. Let $x(s)$ be a characteristic path such that $x'(s) = u(s)$, $x_0 := x(0) = 0$, and $x_S := x(S) = L$. We now derive expressions for x, u, ψ with respect to s.

Substituting the first u in (3.10) with $x'(s)$, we get

$$x'(s)u'(x) = \psi \implies u'(s) = \psi.$$

Multiply (3.11) by $x'(s)$ on both side, we get

$$x'(s)\psi'(x) = x'(s)\rho \implies \psi'(s) = u\rho = J = J_0.$$

Multiply the definition of ψ by $x'(s)$ on both side, we get

$$x'(s)\phi'(x) = x'(s)\psi \implies \phi'(s) = u\psi = uu'(s).$$

Hence we arrive the following system of differential equations:

$$\psi'(s) = J_0 \tag{3.17}$$

$$u'(s) = \psi(s) \tag{3.18}$$

$$x'(s) = u(s) \tag{3.19}$$

$$\phi'(s) = u(s)u'(s) \tag{3.20}$$

and therefore

$$\psi(s) = \psi_0 + J_0 s \tag{3.21}$$

$$u(s) = u_0 + \psi_0 s + \frac{1}{2} J_0 s^2 \tag{3.22}$$

$$x(s) = u_0 s + \frac{1}{2}\psi_0 s^2 + \frac{1}{6} J_0 s^3 \tag{3.23}$$

$$\phi(s) = \frac{1}{2}\left(u^2(s) - u_0^2\right). \tag{3.24}$$

Our goal is to express J_0 as a function of u_0, ϕ_L, L, S and maximize it over S, where S is the ending time of the characteristics. At $s = S$, (3.22), (3.23), (3.24) imply that

$$u_S = u_0 + \psi_0 S + \frac{1}{2} J_0 S^2 \tag{3.25}$$

$$L = u_0 S + \frac{1}{2}\psi_0 S^2 + \frac{1}{6} J_0 S^3 \tag{3.26}$$

$$2\phi_L = u_S^2 - u_0^2. \tag{3.27}$$

Substituting u_S in (3.25) with (3.27), we get

$$\sqrt{2\phi_L + u_0^2} = u_0 + \psi_0 S + \frac{1}{2} J_0 S^2. \tag{3.28}$$

Further substituting $\psi_0 S^2$ in (3.26) with (3.28), we get

$$L = u_0 S + \frac{1}{2}\left(\sqrt{2\phi_L + u_0^2}S - u_0 S - \frac{1}{2}J_0 S^3\right) + \frac{1}{6}J_0 S^3 \tag{3.29}$$

$$= \frac{S}{2}\left(u_0 + \sqrt{2\phi_L + u_0^2}\right) - \frac{1}{12}J_0 S^3, \tag{3.30}$$

and hence

$$J_0(u_0, \phi_L, L, S) = 6\left(u_0 + \sqrt{u_0^2 + 2\phi_L}\right) S^{-2} - 12LS^{-3}. \tag{3.31}$$

Maximizing (3.31) with respect to S yields equation (3.12).

3.1.2 Comparison

A lower bound of this limit can be obtained through solutions from dimensional analysis, which corresponds to the Child' law where the initial velocity is zero. Suppose $\rho = ax^b$, $u = cx^d$, $\psi = ex^f$. By matching the coefficients and the powers in the time-independent Euler-Poisson equations, we get

$$\rho(x) = \frac{2c^2}{9}x^{-\frac{2}{3}} \tag{3.32}$$

$$u(x) = cx^{\frac{2}{3}} \tag{3.33}$$

$$\psi(x) = \frac{2c^2}{3}cx^{\frac{1}{3}} \tag{3.34}$$

for some positive constant c determined by the initial conditions. Note that such solutions exist only when x is positive hence we should assume $x \in [L_1, L_1 + L]$ with $L, L_1 > 0$. According to (3.32), (3.33), (3.34), the flux, initial velocity, and potential difference can

be expressed as

$$u_0 = cL_1^{\frac{2}{3}} \tag{3.35}$$

$$J_0 = \rho u = \frac{2}{9}c^3 \tag{3.36}$$

$$\phi_L = \int_{L_1}^{L_1+L} \psi(x)dx = \frac{c^2}{2}\left((L_1+L)^{\frac{4}{3}} - L_1^{\frac{4}{3}}\right). \tag{3.37}$$

Substituting L_1 in (3.37) with (3.35) yields

$$(2\phi_L + u_0^2)^{\frac{3}{4}} = u_0^{\frac{3}{2}} + c^{\frac{3}{2}}L \tag{3.38}$$

and hence

$$\begin{aligned} J_0 &= \frac{2}{9L^2}\left[(2\phi_L + u_0^2)^{\frac{3}{4}} - u_0^{\frac{3}{2}}\right]^2 \\ &= \frac{2}{9L^2}\left[(2\phi_L + u_0^2)^{\frac{3}{2}} - 2(2\phi_L + u_0^2)^{\frac{3}{4}}u_0^{\frac{3}{2}} + u_0^3\right] \\ &\leq \frac{2}{9L^2}\left[(2\phi_L + u_0^2)^{\frac{3}{2}} + 3(2\phi_L + u_0^2)u_0^{\frac{1}{2}} + 3(2\phi_L + u_0^2)^{\frac{1}{2}}u_0 + u_0^3\right] \\ &= J_{\text{CL}}. \end{aligned} \tag{3.39}$$

In the limit where $u_0 = 0$ (or $L_1 = 0$), $J_0 = J_{\text{CL}}$ and it is called the Child's law.

3.1.3 Limitation

There are two limitations to the CL limit: one due to the algebraic structure of the formula and one due to the validity of characteristics. First of all, since the flux is real, (3.12) implies that

$$u_0^2 + 2\phi_L \geq 0 \tag{3.40}$$

or else J_{CL} becomes imaginary and non-physical. The second limitation is that the characteristic/particle speed u must remain positive in the domain, which is equivalent to

requiring

$$\psi_0 > -\sqrt{2u_0 J_0} \tag{3.41}$$

as explained in [18] and [20]. In figure 3.1 we can see that the characteristics almost become vertical when $\psi_0 \approx -\sqrt{2u_0 J_0}$.

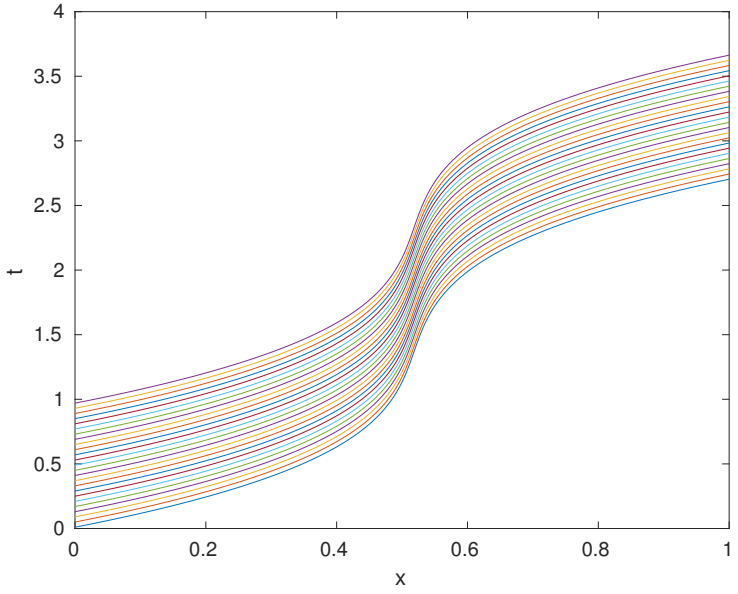

FIGURE 3.1: Characteristics of steady solutions when $\psi_0 \approx -\sqrt{2u_0 J_0}$

3.2 Flux optimization

Among all the admissible sets of boundary conditions (ρ_0, u_0, ϕ_L), we would like to find one that maximizes the flux difference $\bar{J}_{\text{diff}}$. This is equivalent to minimizing the CL limit while keeping $\bar{J}_0$ constant, i.e.

$$\min_{\rho_0, u_0, \phi_L} \bar{J}_{\text{CL}}(u_0, \phi_L, L) \quad \text{s.t. } \bar{J}_0(\rho_0, u_0) = 1 \text{ and } (\rho, u, \phi) \text{ solve } (3.4), (3.5), (3.6) \tag{3.42}$$

which can be solved with Lagrange multipliers.

3.2.1 Lagrange Multiplier

Let

$$\mathcal{L} = \bar{J}_{\text{CL}} + \lambda(\bar{J}_0 - 1) + \int_0^T \int_0^L \nu_1 \cdot (3.4) + \nu_2 \cdot (3.5) + \nu_3 \cdot (3.6) dxdt \tag{3.43}$$

to be the Lagrangian corresponding to our objective function (3.42). Our goal is to solve $\min_{\rho_0, u_0, \phi_L} \mathcal{L}$. To find the minimizer of $\mathcal{L}$, we need to derive its gradients and the conditions such that

$$\partial_{(\rho,u,\phi)} \mathcal{L} = F(\rho_0, u_0, \phi_L). \tag{3.44}$$

for some explicity function F because we only have control over ρ_0, u_0, ϕ_L. Notation-wise, we define

$$\begin{aligned}
\int f(x,t)|_{x=L} dt &:= \int f(L,t) dt \\
\int f(x,t)|_{x=0}^{x=L} dt &:= \int f(L,t) - f(0,t) dt.
\end{aligned}$$

3.2.2 Density

Differentiating the Lagrangian $\mathcal{L}$ with respect to ρ and applying integration by parts, we get

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial \rho} &= \frac{\partial}{\partial \rho} \bar{J}_{\text{CL}} + \frac{\partial}{\partial \rho} \int \lambda(J_0 - 1) dt + \frac{\partial}{\partial \rho} \int \int \nu_1 \partial_t \rho + \nu_1 \partial_x (\rho u) - \nu_3 \rho dx dt \\
&= 0 + \frac{\partial}{\partial \rho} \left[\int \lambda \rho u|_{x=0} dt + \int \nu_1 \rho|_{t=0}^{t=T} dx + \int \nu_1 \rho u|_{x=0}^{x=L} dt \right] \\
&\quad + \frac{\partial}{\partial \rho} \int \int -\rho \partial_t \nu_1 - \rho u \partial_x \nu_1 - \nu_3 \rho dx dt \\
&= \frac{\partial}{\partial \rho} \left[\int (\lambda - \nu_1) \rho u|_{x=0} dt + \int \nu_1 \rho u|_{x=L} dt + \int \nu_1 \rho|_{t=0}^{t=T} dx \right] \\
&\quad + \int \int -\partial_t \nu_1 - u \partial_x \nu_1 - \nu_3 dx dt \\
&= I_1 + I_2 + I_3 + I_4.
\end{aligned}$$

Here are some conditions that help us eliminate I_2, I_3, I_4.

1. If $\nu_1(L,t) = 0$ for all $t \in [0,T]$, then $I_2 = 0$.

2. If $\nu_1(x,t)$ is periodic in t for all $x \in [0,L]$, then $I_3 = 0$.

3. If $\partial_t \nu_1 + u\partial_x \nu_1 + \nu_3 = 0$ for all $(x,t) \in [0,L] \times [0,T]$, then $I_4 = 0$.

Those conditions will be summarized and restated in section 3.2.5.

3.2.3 Velocity

Differentiating the Lagrangian $\mathcal{L}$ with respect to u and applying integration by parts, we get

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial u} &= \frac{\partial}{\partial u}\bar{J}_{\mathrm{CL}} + \frac{\partial}{\partial u}\int \lambda(J_0 - 1)dt + \frac{\partial}{\partial u}\int\int \nu_1 \partial_x(\rho u) + \nu_2 \partial_t u - \nu_2 u \partial_x u dx dt \\
&= \frac{\partial}{\partial u}\left[\int J_{\mathrm{CL}} + \lambda \rho u|_{x=0} dt + \int \nu_1 \rho u|_{x=0}^{x=L} dt + \int \nu_2 u|_{t=0}^{t=T} dx + \int \nu_2 \frac{1}{2} u^2|_{x=0}^{x=L} dt\right] \\
&\quad + \frac{\partial}{\partial u}\int\int -\rho u \partial_x \nu_1 - u\partial_t \nu_2 - \frac{1}{2}u^2 \partial_x \nu_2 dx dt \\
&= \frac{\partial}{\partial u}\left[\int (J_{\mathrm{CL}} + \lambda \rho u - \nu_1 \rho u - \frac{1}{2}\nu_2 u^2)|_{x=0} dt + \int (\nu_1 \rho u + \frac{1}{2}\nu_2 u^2)|_{x=L} dt\right] \\
&\quad + \frac{\partial}{\partial u}\int \nu_2 u|_{t=0}^{t=T} dx + \int\int -\rho \partial_x \nu_1 - \partial_t \nu_2 - u\partial_x \nu_2 dx dt \\
&= I_1 + I_2 + I_3 + I_4.
\end{aligned}$$

Here are some conditions that help us eliminate I_2, I_3, I_4.

1. If $\nu_1(L,t) = \nu_2(L,t) = 0$ for all $t \in [0,T]$, then $I_2 = 0$.

2. If $\nu_2(x,t)$ is periodic in t for all $x \in [0,L]$, then $I_3 = 0$.

3. If $\rho\partial_x \nu_1 + \partial_t \nu_2 + u\partial_x \nu_2 = 0$ for all $(x,t) \in [0,L] \times [0,T]$, then $I_4 = 0$.

Those conditions will be summarized and restated in section 3.2.5.

3.2.4 Potential

Differentiating the Lagrangian $\mathcal{L}$ with respect to ψ and applying integration by parts, we get

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial \psi} &= \frac{\partial}{\partial \psi}\bar{J}_{\mathrm{CL}} + \frac{\partial}{\partial \psi}\int\int -\nu_2\psi + \nu_3\partial_x\psi dxdt \\
&= \frac{\partial}{\partial \psi}\int \frac{2}{9L^2}\left(u_0 + \sqrt{u_0^2 + 2\int \psi dx}\right)^3 dt + \frac{\partial}{\partial \psi}\int\int -\nu_2\psi + \nu_3\partial_x\psi dxdt \\
&= \frac{\partial}{\partial \psi}\int \nu_3\psi|_{x=0}^{x=L}dt + \int\int \frac{2}{3L^2}\frac{\left(u_0 + \sqrt{u_0^2 + 2\int \psi dx}\right)^2}{\sqrt{u_0^2 + 2\int \psi dx}} - \nu_2 - \partial_x\nu_3 dxdt \\
&= I_1 + I_2.
\end{aligned}$$

If $\nu_3(L,t) = 0$ for all $t \in [0,T]$, then $I_2 = \frac{\partial}{\partial \psi}\int -\nu_3\psi|_{x=0}dt$.

3.2.5 Dual Equations

From section 3.2.3, 3.2.2, 3.2.4, we arrive at the dual equations

$$u\partial_x\nu_1 + \partial_t\nu_1 + \nu_3 = 0 \tag{3.45}$$

$$u\partial_x\nu_2 + \rho\partial_x\nu_1 + \partial_t\nu_2 = 0 \tag{3.46}$$

$$\partial_x\nu_3 + \nu_2 - g_1 = 0 \tag{3.47}$$

(for g_1 defined in (3.54)) with zero boundary conditions on $x = L$

$$\nu_i(L,t) = 0 \tag{3.48}$$

and periodic boundary conditions in t

$$\nu_i(x,0) = \nu_i(x,T) \tag{3.49}$$

for $i = 1, 2, 3$. If (3.45), (3.46), (3.47), (3.48), (3.49) are all satisfied, then the gradient of $\mathcal{L}$ is given by

$$\partial_\rho \mathcal{L} = (\lambda - \nu_1(0,t))u_0 \tag{3.50}$$

$$\partial_u \mathcal{L} = \lambda\rho_0 - \nu_1(0,t)\rho_0 - \nu_2(0,t)u_0 + g_2 \tag{3.51}$$

$$\partial_\psi \mathcal{L} = -\nu_3(0,t) \tag{3.52}$$

$$\partial_\lambda \mathcal{L} = \bar{J}_0 - 1 \tag{3.53}$$

where

$$g_1 = \frac{2}{3L^2} \frac{\left(u_0 + \sqrt{u_0^2 + 2\int \psi dx}\right)^2}{\sqrt{u_0^2 + 2\int \psi dx}} \tag{3.54}$$

$$g_2 = \frac{2}{3L^2} \left(u_0 + \sqrt{u_0^2 + 2\int \psi dx}\right)^2 \left(1 + \frac{u_0}{\sqrt{u_0^2 + 2\int \psi dx}}\right). \tag{3.55}$$

We expect that performing gradient descent in such direction, once converges, will yield a local minimum of $\bar{J}_{\text{CL}} < 1 = \bar{J}_0$.

3.3 Algorithm

Denote

$$\xi(x,t) := (\rho(x,t), u(x,t), \psi(x,t), \lambda)$$

$$\zeta(x,t) := (\nu_1(x,t), \nu_2(x,t), \nu_3(x,t))$$

to be the primal and dual variables respectively. We apply gradient descent on the Lagrangian $\mathcal{L}$ by updating the primal and dual variables according to algorithm 3.

Algorithm 3 Flux Limit Variation Principle

while $||\nabla_{\xi(0,t)}\mathcal{L}||_2 >$tol **do**

Solve $\xi(x,t)$ forward in x with $\xi(0,t)$ and (3.4), (3.5), (3.6).

Solve $\zeta(x,t)$ backward in x with $\zeta(L,t)=0$ and (3.45), (3.46), (3.47).

Update $\xi(0,t) \longleftarrow \xi(0,t) - \epsilon\nabla_{\xi(0,t)}\mathcal{L}$ with (3.50), (3.51), (3.52), (3.53).

end while

For numerical purpose, it is advantageous to put (3.4), (3.5), (3.6) into conservative forms by defining new variables $\eta_1 = \rho u, \eta_2 = u^2, \psi = \partial_x\phi$ so that

$$\partial_x\eta_1 + \partial_t\left(\frac{\eta_1}{\sqrt{\eta_2}}\right) = 0 \tag{3.56}$$

$$\partial_x\eta_2 + 2\partial_t(\sqrt{\eta_2}) - 2\psi = 0 \tag{3.57}$$

$$\partial_x\psi - \frac{\eta_1}{\sqrt{\eta_2}} = 0. \tag{3.58}$$

Assuming that η_2 does not vanish, equations (3.56), (3.57), (3.58) can be solved stably and consistently using upwind method, which will be our choice of numerical method for both the forward equations and backward equations.

There are three major stopping criteria for our method:

1. $\mathcal{L}$ converges: local minimum is obtained
2. ρ or u becomes negative: characteristics/numerical method breaks down
3. $\bar{J}_{\mathrm{CL}}$ becomes imaginary: condition (3.40) is violated

Depending on the initialization, the algorithm might stop due to different reason. Although we often observe the last two cases before the first occurs, those solutions are nontrivial and do exceed CL limit.

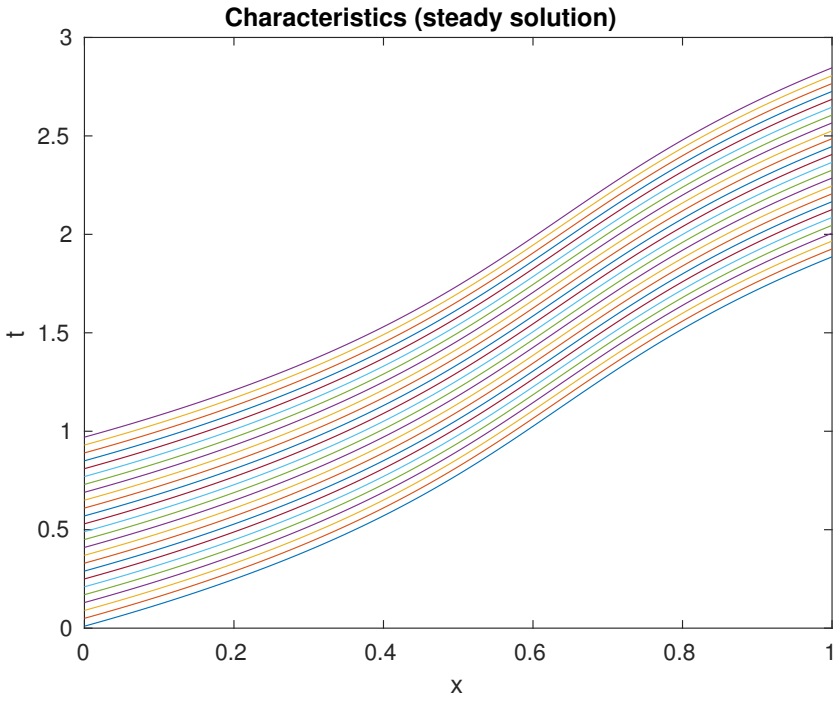

FIGURE 3.2: Steady solution with parallel characteristics

3.4 Numerical Results

In this section we will show and analyze some numerical results. In general, When the initialization is time-homogeneous, the resulting current always converges to the CL limit. When the initialization is not time-homogeneous, the resulting current converges to/breaks down at different values depending on the initialization.

3.4.1 Steady Solution

When the initialization is time-homogeneous, i.e. the boundary condition is a constant function, the resulting current $\bar{J}$ always converges to the CL limit $\bar{J}_{\text{CL}}$ as expected. The characteristics of our solution are parallel and no shock is formed, as shown in figure 3.2.

3.4.2 Break Down of Characteristics

When the initialization is not time-homogeneous, the algorithm often breaks down because either the velocity becomes negative, where our numerical PDE solver for (3.56),

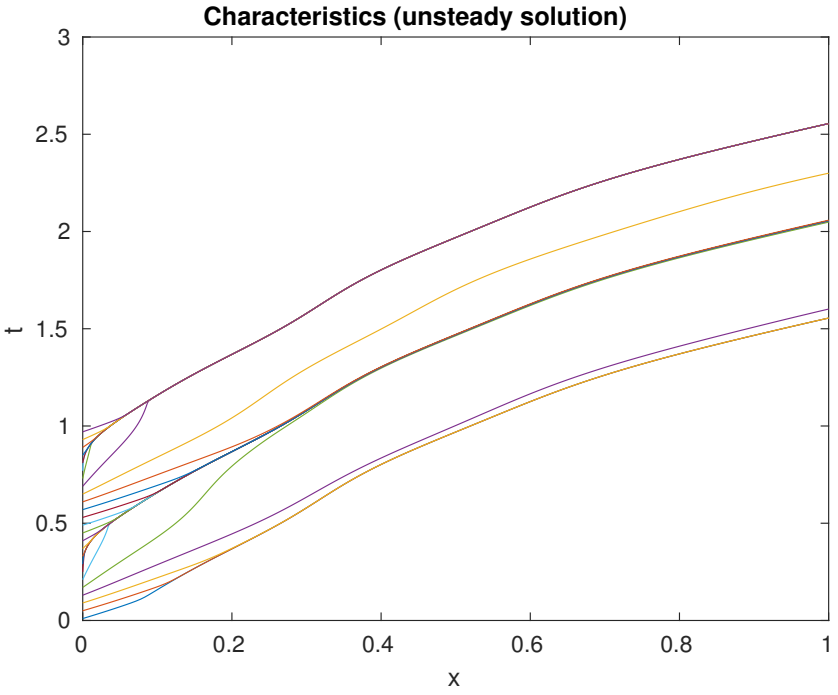

FIGURE 3.3: Formation of shocks and nearly vertical characteristics

(3.57), (3.58) is no longer valid, or when the density becomes negative, in which the solution is no longer physical.

In one of our experiment, $\bar{J}$ exceeds $1.25 \cdot \bar{J}_{\text{CL}}$ before the velocity of our solution becomes negative at some point in the interior of the domain. Its characteristics are shown in figure 3.3.

3.4.3 Break Down of CL Limit

Another reason for our algorithm to break is the violation of condition (3.40), where J_{CL} is no longer real and hence non-physical. This is more of a constraint to our algebraic formulation rather than a physical constraints.

In one of our experiment, $\bar{J}$ exceeds $1.02 \cdot \bar{J}_{\text{CL}}$ before J_{CL} becomes complex. There is no indication of shocks in its characteristics, but we can see that its characteristics are concentrated towards certain region, as shown in figure 3.3.

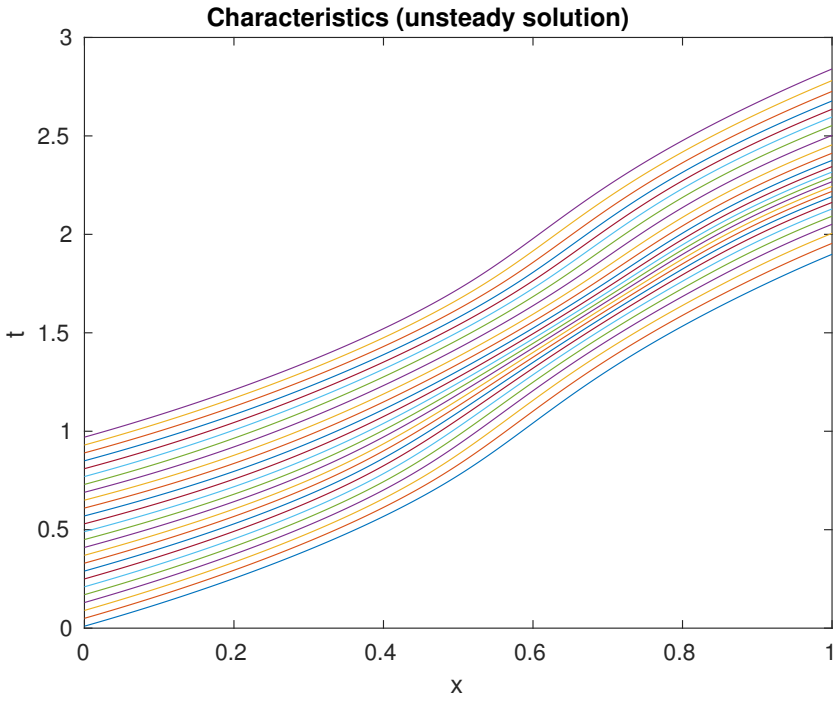

FIGURE 3.4: Nonuniform density in characteristics

3.4.4 Performance

We search over the parameter space $c_i \in (0, 1)$, where

$$\rho_0 = 1 - c_1 \cos(t)$$

$$u_0 = 1$$

$$\psi_0 = c_0 + c_1 \sin(t) - c_2 \cos(t)$$

and managed to find solutions whose current is around 40% above the CL limit. However, all the optimal solutions we found, except the steady solutions, are at the boundary of our solution domain.

3.5 Conclusion

In this chapter we re-derive, analyze, and improve the flux limit of Euler-Poisson equations through the method of characteristics and variation principle. The result from our

Lagrangian optimization is consistent with the time-independent CL limit. On the other hand, most of our time-dependent optimizers exceed CL limit but eventually hit the boundary of our solution space before they converge to a local minimum. The improvement ranges from 2% to 40% depending on the initialization, and we would like to show analytically that 40% is the flux limit of time-dependent Euler-Poisson equations.

Appendix A

Signal Segmentation

A.1 Poisson Summation Formula

Define continuous Dirac comb $C_\Delta : \mathbb{R} \to \mathbb{R}$ to be

$$C_\Delta(t) = \sum_{n \in \mathbb{Z}} \delta_{n\Delta}(t). \tag{A.1}$$

where $\Delta \neq 0$ and δ is the Dirac delta distribution.

Proposition A.1. *(Poisson Summation Formula)[21]*

The Dirac comb C_Δ has Fourier transform (in the sense of distributions) given by

$$\hat{C}_\Delta = \frac{2\pi}{\Delta} C_{2\pi/\Delta}. \tag{A.2}$$

A.2 Proof of Corollary 1.8

Without loss of generality, let

$$f(t) = \sin(t)$$
$$g(t) = \frac{1}{N} \sum_{n=-\infty}^{\infty} f(n\Delta)\psi(t - n\Delta)$$

and denote

$$\mathcal{F}_p(f)(k) := \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-ikt} dt$$

to be the Fourier transform of function with period 2π. It suffices to show that $\mathcal{F}_p(f) = \mathcal{F}_p(g)$. For simplicity denote $C = \frac{N}{\sin(\Delta)}$. By definition,

$$
\begin{aligned}
\mathcal{F}_p(f)(k) &= \int_0^{2\pi} \frac{1}{4\pi i}(e^{it} - e^{-it})e^{-ikt}dt \\
&= \frac{-1}{4\pi(1-k)}(e^{-2\pi ik} - 1) - \frac{1}{4\pi(1+k)}(e^{-2\pi ik} - 1) \\
&= \frac{1}{2\pi(1-k^2)}(1 - e^{-2\pi ik}) \\
\mathcal{F}_p(\psi)(k) &= \frac{C}{2\pi}\int_{-\Delta}^{0} \sin(\Delta + t)e^{-ikt}dt + \frac{C}{2\pi}\int_0^{\Delta} \sin(\Delta - t)e^{-ikt}dt \\
&= \frac{C}{4\pi i}\int_{-\Delta}^{0} e^{i(\Delta+t-kt)} - e^{-i(\Delta+t+kt)}dt \\
&\quad + \frac{C}{4\pi i}\int_0^{\Delta} e^{i(\Delta-t-kt)} - e^{-i(\Delta-t+kt)}dt \\
&= \frac{C}{4\pi i}\left[\frac{e^{i\Delta} - e^{ik\Delta}}{i(1-k)} + \frac{e^{-i\Delta} - e^{-ik\Delta}}{i(1+k)}\right] \\
&\quad + \frac{C}{4\pi i}\left[\frac{e^{-ik\Delta} - e^{i\Delta}}{i(-1-k)} + \frac{e^{ik\Delta} - e^{-i\Delta}}{i(-1+k)}\right] \\
&= \frac{C}{2\pi(1-k^2)}\left(e^{ik\Delta} + e^{-ik\Delta} - e^{i\Delta} - e^{-i\Delta}\right) \\
\mathcal{F}_p(g)(k) &= \frac{1}{2\pi}\int_0^{2\pi} \frac{1}{N}\sum_{n=0}^{N-1} f(n\Delta)\psi(t - n\Delta)e^{-ikt}dt \\
&= \frac{1}{N}\sum_{n=0}^{N-1} f(n\Delta)\mathcal{F}_p(\psi)(k)e^{-ikn\Delta} \\
&= \mathcal{F}_p(\psi)(k)\frac{1}{2Ni}\sum_{n=0}^{N-1}(e^{in\Delta} - e^{-in\Delta})e^{-ikn\Delta} \\
&= \mathcal{F}_p(\psi)(k)\frac{1}{2Ni}\left(\frac{1 - e^{-2\pi ik}}{1 - e^{i(1-k)\Delta}} - \frac{1 - e^{-2\pi ik}}{1 - e^{-i(1+k)\Delta}}\right) \\
&= \frac{C}{2\pi(1-k^2)}\left(e^{ik\Delta} + e^{-ik\Delta} - e^{i\Delta} - e^{-i\Delta}\right) \\
&\quad \cdot \frac{1}{2Ni}\left(\frac{e^{i\Delta} - e^{-i\Delta}}{e^{ik\Delta} + e^{-ik\Delta} - e^{i\Delta} - e^{-i\Delta}}\right)(1 - e^{-2\pi ik}) \\
&= \frac{1}{2\pi(1-k^2)}(e^{-2\pi ik} - 1) \cdot \frac{C\sin(\Delta)}{N} \\
&= \mathcal{F}_p(f)(k)
\end{aligned}
$$

A.3 Analysis of kernel η

To write $\phi = \psi * \eta$, we will calculate the Fourier transform of ϕ and ψ. We will now compute the $\hat{\phi}$ and $\hat{\eta}$. By symmetry, we define

$$v(k) := \int_0^{\Delta} e^{-itk}\psi(t)dt \tag{A.3}$$

$$z(k) := \int_0^{2\Delta} e^{-itk}\phi(t)dt \tag{A.4}$$

so that

$$\hat{\psi}(k) = v(k) + v(-k) \tag{A.5}$$

$$\hat{\phi}(k) = z(k) + z(-k). \tag{A.6}$$

For simplicity, we will denote C to be the normalizing constant which will be specified in the end. Express the Fourier transform of (1.30) with (A.3), we get

$$\begin{aligned} Cz(k) &= 4\cos(\Delta)e^{-i\Delta k}v(-k) + (4\cos^2(\Delta) + 2)v(k) \\ &\quad + (4\cos^2(\Delta) - 2)e^{-2i\Delta k}v(-k) + 8\cos^3(\Delta)e^{-i\Delta k}v(k) \\ &= \left(2 + 4\cos^2(\Delta) + 8\cos^3(\Delta)e^{-i\Delta k}\right)v(k) \\ &\quad + \left(4\cos(\Delta)e^{-i\Delta k} + 4\cos^2(\Delta)e^{-2i\Delta k} - 2e^{-2i\Delta k}\right)v(-k) \end{aligned}$$

and hence

$$
\begin{aligned}
C\hat{\phi}(k) &= Cz(k) + Cz(-k) \\
&= \begin{bmatrix} v(k)+v(-k) & v(k)-v(-k) \end{bmatrix} \cdot \\
&\begin{bmatrix} 2 + 4\cos^2(\Delta) + (4\cos(\Delta) + 8\cos^3(\Delta))\cos(\Delta k) + (4\cos^2(\Delta) - 2)\cos(2\Delta k) \\ i\,(4\cos(\Delta)\sin(\Delta k) + (4\cos^2(\Delta) - 2)\sin(2\Delta k) - 8\cos^3(\Delta)\sin(\Delta k)) \end{bmatrix} \\
&= \begin{bmatrix} v(k)+v(-k) & v(k)-v(-k) \end{bmatrix} \cdot \\
&\begin{bmatrix} 4 + (4\cos(\Delta) + 8\cos^3(\Delta))\cos(\Delta k) + (8\cos^2(\Delta) - 4)\cos^2(\Delta k) \\ i(4\cos(\Delta) - 8\cos^3(\Delta) + 8\cos^2(\Delta)\cos(\Delta k) - 4\cos(\Delta k))\sin(\Delta k) \end{bmatrix} \\
&= \frac{8}{k^2-1} \begin{bmatrix} \cos(\Delta) - \cos(\Delta k) & k\sin(\Delta) - \sin(\Delta k) \end{bmatrix} \cdot \\
&\begin{bmatrix} 1 + \cos(\Delta)(1 + 2\cos^2(\Delta))\cos(\Delta k) + \cos(2\Delta)\cos^2(\Delta k) \\ -\cos(2\Delta)(\cos(\Delta) - \cos(\Delta k))\sin(\Delta k) \end{bmatrix} \\
&= \frac{8}{1-k^2}(\cos(\Delta) - \cos(\Delta k)) \\
&\quad \cdot [1 + \cos(2\Delta) + \cos(\Delta)(1 + 2\cos^2(\Delta))\cos(\Delta k) - k\sin(\Delta)\cos(2\Delta)\sin(\Delta k)] \\
&= C\hat{\psi} \cdot \hat{\eta}
\end{aligned}
$$

where

$$
\begin{aligned}
\hat{\eta}(k) := {} & 4c\sin^2(\Delta)[1 + \cos(2\Delta) + \cos(\Delta)(1 + 2\cos^2(\Delta))\cos(\Delta k)] \\
& - 4c\sin^2(\Delta)[k\sin(\Delta)\cos(2\Delta)\sin(\Delta k)]
\end{aligned} \tag{A.7}
$$

c being the same as in (1.30). According to (A.7) in, η has inverse Fourier transform and equals to

$$
\eta = c_0 + c_1(\theta_\Delta + \theta_{-\Delta}) - c_2(\theta_\Delta - \theta_{-\Delta})D \tag{A.8}
$$

where $\theta_\Delta(u)(t) := u(t-\Delta)$ is the shift operator and $Du(t) := u'(t)$ is the differential operator.

Appendix B

MRA

B.1 Spectrum and Bi-spectrum Calculation

Since

$$\begin{aligned}
\mathbb{E}\left[\hat{\xi}(k_1)\hat{\xi}(k_2)\right] &= \frac{1}{N^2}\mathbb{E}\left[\left(\sum_{n=1}^{N}\xi(n)e^{-ink_1\Delta}\right)\left(\sum_{n=1}^{N}\xi(n)e^{-ink_2\Delta}\right)\right] \\
&= \frac{1}{N^2}\mathbb{E}\left[\sum_{n=1}^{N}\xi^2(n)e^{-in(k_1+k_2)\Delta} + \sum_{n\neq m}^{N-1}\xi(n)\xi(m)e^{-i(nk_1+mk_2)\Delta}\right] \\
&= \frac{1}{N^2}\sum_{n=1}^{N}\mathbb{E}\left[\xi^2(n)\right]e^{-in(k_1+k_2)\Delta} + 0 \\
&= \frac{1}{N}\delta_0(k_1+k_2) \qquad \text{(B.1)}
\end{aligned}$$

direct computation implies that

$$\begin{aligned}
\mathbb{E}\left[S_y(k)\right] &:= \mathbb{E}\left[\hat{y}(k)\hat{y}(-k)\right] \\
&= \mathbb{E}\left[\left(e^{igk\Delta}\hat{x}(k) + \sigma\hat{\xi}(k)\right)\left(e^{-igk\Delta}\hat{x}(-k) + \sigma\hat{\xi}(-k)\right)\right] \\
&= \hat{x}(k)\hat{x}(-k) + \sigma^2\mathbb{E}\left[\hat{\xi}(k)\hat{\xi}(-k)\right] \\
&\quad + \sigma e^{igk\Delta}\hat{x}(k)\mathbb{E}\left[\hat{\xi}(-k)\right] + \sigma e^{-igk\Delta}\hat{x}(k)\mathbb{E}\left[\hat{\xi}(k)\right] \\
&= S_x(k) + \frac{\sigma^2}{N}
\end{aligned} \tag{B.2}$$

and

$$\begin{aligned}
\mathbb{E}[B_y(k_1,k_2)] &:= \mathbb{E}\left[\hat{y}(k_1)\hat{y}(k_2)\hat{y}(-k_1-k_2)\right] \\
&= \hat{x}(k_1)\hat{x}(k_2)\hat{x}(-k_1-k_2) \\
&\quad + \sigma e^{-igk_1\Delta}\hat{x}(k_2)\hat{x}(-k_1-k_2)\mathbb{E}\left[\hat{\xi}(k_1)\right] \\
&\quad + \sigma e^{-igk_2\Delta}\hat{x}(k_1)\hat{x}(-k_1-k_2)\mathbb{E}\left[\hat{\xi}(k_2)\right] \\
&\quad + \sigma e^{ig(k_1+k_2)\Delta}\hat{x}(k_1)\hat{x}(k_2)\mathbb{E}\left[\hat{\xi}(-k_1-k_2)\right] \\
&\quad + \sigma^2 e^{igk_1\Delta}\hat{x}(k_1)\mathbb{E}\left[\hat{\xi}(k_2)\hat{\xi}(-k_1-k_2)\right] \\
&\quad + \sigma^2 e^{igk_2\Delta}\hat{x}(k_2)\mathbb{E}\left[\hat{\xi}(k_1)\hat{\xi}(-k_1-k_2)\right] \\
&\quad + \sigma^2 e^{-ig(k_1+k_2)\Delta}\hat{x}(-k_1-k_2)\mathbb{E}\left[\hat{\xi}(k_1)\hat{\xi}(k_2)\right] \\
&\quad + \sigma^3\mathbb{E}\left[\hat{\xi}(k_1)\hat{\xi}(k_2)\hat{\xi}(-k_1-k_2)\right] \\
&= B_x(k_1,k_2) + \frac{\sigma^3}{N^2} \\
&\quad + \frac{\sigma^2}{N}\left[e^{igk_1\Delta}\delta_0(-k_1) + e^{igk_2\Delta}\delta_0(-k_2) + e^{-ig(k_1+k_2)}\delta_0(k_1+k_2)\right] \\
&= B_x(k_1,k_2) + \frac{\sigma^3}{N^2} \quad (\text{for } k_1, k_2, k_1+k_2 \neq 0)
\end{aligned} \tag{B.3}$$

B.2 Cyclotomic Polynomial

Definition B.1. Let $N \in \mathbb{N}$. The N-th cyclotomic polynomialis defined as

$$\Phi_L(x) := \prod_{\substack{1 \le n < N, \\ \gcd(n,N)=1}} (x - e^{\frac{2\pi i n}{N}}) \tag{B.4}$$

Proposition B.2. *For primes $N \geq 3$, the Nth cyclotomic polynomials are irreducible over the field of the rational numbers.*

Proof. Although the claim is true for $N \geq 3$ in general, it suffices to prove the case of prime for our purpose. Recall Eisenstein's criterion: For polynomials of the form $Q(x) = \sum_{n=0}^{N} a_n x^n$, if there exists prime p such that $p \mid a_n$ for all $n < N$, $p \nmid a_L$, and $p^2 \nmid a_0$, then Q is irreducible over $\mathbb{Q}$. Since

$$Q(x) = \frac{(x+1)^N - 1}{x} = \sum_{n=0}^{N-1} \binom{N}{n+1} x^n \tag{B.5}$$

satisfies Eisenstein's criterion, it is irreducible. Substituting $y = x + 1$, we can deduce that

$$\Phi_N(x) := \sum_{n=0}^{N-1} y^n = \frac{y^N - 1}{y - 1} \tag{B.6}$$

is also irreducible. □

B.3 Ratio Estimator

Consider two random variable X, Y with mean μ_X, μ_Y; we want to estimate $\mu_X \mu_Y^{-1}$ with $\mathbb{E}\left[XY^{-1}\right]$ despite the fact that in general $\mathbb{E}\left[XY^{-1}\right] \neq \mu_X \mu_Y^{-1}$. In this section we use Taylor expansion to analyze the expectation and variance of our estimator and its relation to $\mu_X \mu_Y^{-1}$. Let $f(X, Y) = XY^{-1}$, $\theta = (\mu_X, \mu_Y)$, $\Delta_X = X - \mu_X$, and $\Delta_Y = Y - \mu_Y$. The

second order Taylor expansion in $\mathbb{E}[f(X,Y)]$ is

$$\begin{aligned}\mathbb{E}\left[f(X,Y)\right] &\approx \mathbb{E}\left[f(\theta) + f_x(\theta)\Delta_X + f_y(\theta)\Delta_Y\right] \\ &\quad + \frac{1}{2}\mathbb{E}\left[f_{xx}(\theta)\Delta_X^2 + f_{yy}(\theta)\Delta_Y^2 + 2f_{xy}(\theta)\Delta_X\Delta_Y\right] \\ &= f(\theta) + \frac{1}{2}f_{xx}(\theta)\mathrm{Var}[X] + \frac{1}{2}f_{yy}(\theta)\mathrm{Var}[Y] + f_{xy}(\theta)\mathrm{Cov}[X,Y] \\ &= \mu_X\mu_Y^{-1} + \mu_X\mu_Y^{-3}\sigma_Y^2 - \mu_Y^{-2}\mathrm{Cov}[X,Y] \end{aligned} \tag{B.7}$$

and the first order Taylor expansion in $\mathbb{E}[f(X,Y)^2]$ is

$$\begin{aligned}\mathbb{E}\left[f(X,Y)^2\right] &\approx \mathbb{E}\left[(f(\theta) + f_x(\theta)\Delta_X + f_y(\theta)\Delta_Y)^2\right] \\ &= f(\theta)^2 + f_x(\theta)^2\mathrm{Var}[X] + f_y(\theta)^2\mathrm{Var}[Y] + 2f_x(\theta)f_y(\theta)\mathrm{Cov}[X,Y] \\ &= \mu_X^2\mu_Y^{-2} + \mu_Y^{-2}\sigma_X^2 + \mu_X^2\mu_Y^{-4}\sigma_Y^2 - 2\mu_X\mu_Y^{-3}\mathrm{Cov}[X,Y]. \end{aligned} \tag{B.8}$$

Therefore the variance of the ratio estimator is approximately

$$\mathrm{Var}\left[f(X,Y)\right] \approx \mu_Y^{-2}\sigma_X^2 + \mu_X^2\mu_Y^{-4}\sigma_Y^2 - 2\mu_X\mu_Y^{-3}\mathrm{Cov}[X,Y]. \tag{B.9}$$

B.4 Fermats Little Theorem

Corollary B.3. *If N is a prime, then $2^N \equiv 2 (mod\ N)$*

Proof. Let $z = e^{i\Delta}$. Expand the 2^N terms in p in the form of $Q(z) = \sum_{n=0}^{N-1} a_n z^n$. Since $Q(z) - 2 = 0$, by the irreducibly of the Nth cyclotomic polynomials, $Q(z) - 2$ must be a multiple of the N-th cyclotomic polynomials $\sum_{n=0}^{N-1} z^n$. Hence $2^N - 2 \equiv 0 \,(\mathrm{mod}\ N)$. □

Bibliography

[1] Michael Unser. Sampling-50 years after shannon. *Proceedings of the IEEE*, 88:569 – 587, 05 2000. doi: 10.1109/5.843002.

[2] Jonathan Siegel Russel Caflisch, Hung Hsu Chou. Effcient frame projection of amplitude-modulated signals. *Computational Applied Mathematics Publications*, 2017.

[3] C. E. Shannon. Communication in the presence of noise. *Proceedings of the IRE*, 37 (1):10–21, Jan 1949. ISSN 0096-8390. doi: 10.1109/JRPROC.1949.232969.

[4] Carl De Boor, Ronald A DeVore, and Amos Ron. Approximation from shift-invariant subspaces of $l_2(\mathbb{R}^d)$. *Transactions of the American Mathematical Society*, 341(2):787–806, 1994.

[5] Ole Christensen. *Frames and bases: An introductory course*. Springer Science & Business Media, 2008.

[6] Jonathan Siegel. Shift invariant subspaces and applications to signal fragmentation. *Computational Applied Mathematics Publications*, 2017.

[7] A. Singer and Y. Shkolnisky. Three-dimensional structure determination from common lines in cryo-em by eigenvectors and semidefinite programming. *SIAM Journal on Imaging Sciences*, 4(2):543–572, 2011. doi: 10.1137/090767777. URL https://doi.org/10.1137/090767777.

[8] Amit Singer. Angular Synchronization by Eigenvectors and Semidefinite Programming. *arXiv e-prints*, art. arXiv:0905.3174, May 2009.

[9] Cecilia Aguerrebere, Mauricio Delbracio, Alberto Bartesaghi, and Guillermo Sapiro. Fundamental limits in multi-image alignment. *CoRR*, abs/1602.01541, 2016. URL http://arxiv.org/abs/1602.01541.

[10] Alex Barnett, Leslie Greengard, Andras Pataki, and Marina Spivak. Rapid solution of the cryo-EM reconstruction problem by frequency marching. *arXiv e-prints*, art. arXiv:1610.00404, Oct 2016.

[11] Afonso S. Bandeira, Ben Blum-Smith, Joe Kileel, Amelia Perry, Jonathan Weed, and Alexander S. Wein. Estimation under group actions: recovering orbits from invariants. *arXiv e-prints*, art. arXiv:1712.10163, Dec 2017.

[12] Afonso S. Bandeira, Philippe Rigollet, and Jonathan Weed. Optimal rates of estimation for multi-reference alignment. *arXiv e-prints*, art. arXiv:1702.08546, Feb 2017.

[13] Amelia Perry, Jonathan Weed, Afonso S. Bandeira, Philippe Rigollet, and Amit Singer. The sample complexity of multi-reference alignment. *CoRR*, abs/1707.00943, 2017. URL http://arxiv.org/abs/1707.00943.

[14] Leonid I. Rudin, Stanley Osher, and Emad Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60(1):259 – 268, 1992. ISSN 0167-2789. doi: https://doi.org/10.1016/0167-2789(92)90242-F. URL http://www.sciencedirect.com/science/article/pii/016727899290242F.

[15] T. Goldstein and S. Osher. The split bregman method for l1-regularized problems. *SIAM Journal on Imaging Sciences*, 2(2):323–343, 2009. doi: 10.1137/080725891. URL https://doi.org/10.1137/080725891.

[16] C. D. Child. Discharge from hot cao. *Phys. Rev. (Series I)*, 32:492–511, May 1911. doi: 10.1103/PhysRevSeriesI.32.492. URL https://link.aps.org/doi/10.1103/PhysRevSeriesI.32.492.

[17] Irving Langmuir. The effect of space charge and residual gases on thermionic currents in high vacuum. *Phys. Rev.*, 2:450–486, Dec 1913. doi: 10.1103/PhysRev.2.450. URL https://link.aps.org/doi/10.1103/PhysRev.2.450.

[18] George Jaffé. On the currents carried by electrons of uniform initial velocity. *Phys. Rev.*, 65:91–98, Feb 1944. doi: 10.1103/PhysRev.65.91. URL https://link.aps.org/doi/10.1103/PhysRev.65.91.

[19] P V. Akimov, H Schamel, Heidrun Kolinsky, A Ender, and V Kuznetsov. The true nature of space-charge-limited currents in electron vacuum diodes: A lagrangian revision with corrections. *Physics of Plasmas*, 8:3788–3798, 08 2001. doi: 10.1063/1.1383287.

[20] R. E. Caflisch and M. S. Rosin. Beyond the child-langmuir limit. *Phys. Rev. E*, 85: 056408, May 2012. doi: 10.1103/PhysRevE.85.056408. URL https://link.aps.org/doi/10.1103/PhysRevE.85.056408.

[21] Antoni Zygmund. *Trigonometric series*. Cambridge Mathematical Library. Cambridge University, 1968. ISBN 9780521358859.