On Numerics and Inverse Problems

Peter Caday

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2015

Reading Committee:

Gunther Uhlmann, Chair

Hart Smith

Ken Bube

Program Authorized to Offer Degree:
Department of Mathematics

University of Washington

Abstract

On Numerics and Inverse Problems

Peter Caday

Chair of the Supervisory Committee:
Prof. Gunther Uhlmann
Mathematics

In this thesis, two projects in inverse problems are described. The first concerns a simple mathematical model of synthetic aperture radar with undirected beam, modeled as a 2D circular Radon transform with centers restricted to a plane curve γ. From the work of Stefanov and Uhlmann (2013), it is known that this operator is microlocally non-injective locally, but microlocally injective globally if γ is closed. The problem for non-closed γ is examined; it is shown that the Radon transform R_γ is microlocally non-injective to leading order if a certain geometric condition is satisfied. Known examples where this condition holds are given. Numerical simulations demonstrate R_γ's microlocal non-injectivity for a single curve, and for a four-curve setup satisfying the geometric condition.

The second project involves the implementation of an algorithm by de Hoop, Uhlmann, Vasy, and Wendt (2013), with refinements, for computing generic Fourier integral operators (FIOs) associated with canonical graphs, possibly involving caustics. The algorithm can be divided into two parts: a local component that approximately evaluates an FIO $A\colon C_0^\infty(X) \to \mathcal{D}'(Y)$ expressed in the oscillatory integral form $Af(y) = \int e^{i\phi(y,\xi)} a(y,\xi)\, \hat{f}(\xi)\, d\xi$, modulo an error operator of order $\frac{1}{2}$ less than the order of A, and a global component that expresses an arbitrary FIO associated with a canonical graph as a finite sum of these local oscillatory integrals composed with appropriate coordinate changes. A numerical implementation of their algorithm is demonstrated and successfully applied to a variety of FIOs

associated with canonical graphs. This algorithm is designed to be easy-to-use for future researchers and the code is freely available from the author.

TABLE OF CONTENTS

LIST OF FIGURES

thanks

This thesis is written with thanks to my fiancée Lisa for her love and support;

to Mom, Dad, and my sister Katie; and

to my patient advisors Gunther Uhlmann and Plamen Stefanov.

Some of this work was conducted under the much appreciated support of NSF RTG grant 0838212.

Chapter 1

BACKGROUND

We will begin with some background material on microlocal analysis. The standard references for this material is Hörmander's books [2], and other references include Duistermaat [1] and Trèves [4].

Before beginning, we recall some standard multi-index notation. If $\alpha = (\alpha_1, \ldots, \alpha_n)$ is a multi-index (a vector of nonnegative integers), then $|\alpha| = \alpha_1 + \cdots + \alpha_n$, and $\alpha! = \alpha_1!\alpha_2! \cdots \alpha_n!$. The derivative operator ∂_x^α, where x is an n-dimensional variable, means the $|\alpha|$-order operator $\partial_{x_1}^{\alpha_1} \cdots \partial_{x_n}^{\alpha_n}$. Also, x^α means the monomial $x_1^{\alpha_1} \cdots x_n^{\alpha_n}$, or if $x = (x_1, \ldots, x_n)$ is a vector, the same product of its entries. Finally, for multi-indices α, β, let

$$\binom{\alpha}{\beta} = \binom{\alpha_1}{\beta_1} \cdots \binom{\alpha_n}{\beta_n}. \tag{1.0.1}$$

1.1 Wavefront Set

The basic idea of microlocal analysis is to study both the position and direction of the singularities of a distribution[1]. This basic notion is captured by the wavefront set, which is the microlocal analogue of the singular support. Given a distribution $u \in \mathcal{D}'(X)$, the wavefront set $\mathrm{WF}(u) \subseteq T^*X \setminus 0$ is defined by

$$\mathrm{WF}(u) = (T^*X \backslash 0) \backslash \left\{ (x_0, \xi_0) \;\middle|\; \exists \epsilon, U \ni x : \forall \rho \in C_0^\infty(U),\, \widehat{\rho u}(t\xi) = O(t^{-N})\, \forall \xi : \left| \frac{\xi}{|\xi|} - \frac{\xi_0}{|\xi|_0} \right| < \epsilon \, \forall N \right\}.$$

In other words, (x, ξ) belongs to the wavefront set if after localizing u sufficiently close to x, its Fourier transform decays superalgebraically (faster than any polynomial) in a neighborhood of ξ. The basis of this definition is that the Fourier transform of a smooth function decays

[1] Here and throughout "singularity" refers to a distribution failing to be C^∞.

superalgebraically, since derivatives in the spatial domain correspond to multiplication by monomials in the Fourier domain (and a power of i). By extension, it is reasonable to think of u as being smooth in a direction ξ if $\hat{u}(\xi)$ decays similarly in a conic neighborhood of that direction; to provide spatial locality we multiply by a cutoff ρ.

It is easy to check that $\mathrm{WF}(u)$ projects to the singular support of u:

Proposition 1.1. *Let $\pi\colon T^*X \to X$ be the canonical projection. Then* $\operatorname{singsupp} u = \pi(\mathrm{WF}(u))$.

Proof. Let $x \in X$. First, suppose $x \notin \pi(\mathrm{WF}(u))$. Then for all $\xi \in S^*X$, there exist U_ξ, ϵ_ξ satisfying the conditions in (1.1). Since S^*X is compact it can be covered by finitely many of the sets $\{|\xi' - \xi| < \epsilon_\xi$, associated with $\xi_1, \dots, \xi_m$, say. Taking $U = U_{\xi_1} \cap \cdots \cap U_{\xi_m}$, we have $\widehat{\rho u}(t\xi) = O(|\xi|^{-N})$ for all N for $\rho \in C_0^\infty(U)$, implying $x \notin \operatorname{singsupp} u$.

Conversely, if $x \notin \operatorname{singsupp} u$, there exists $\rho \in C_0^\infty(X)$ nonzero at x such that $\rho u \in C^\infty$. Then $\widehat{\rho u}$ decays superalgebraically in all directions, implying that $(x, \xi) \notin \mathrm{WF}(u)$ for all ξ. □

As an example of the directionality of the wavefront set, we can consider an indicator function of a region Ω. Here the wavefront set consists of (x, ξ) which are conormal to the boundary.

Example. Let $\Omega \subseteq \mathbb{R}^n$ be a smooth submanifold, with defining function ρ (so $\Omega = \{\rho \leq 0\}$, $d\rho|_{\partial\Omega} \neq 0$). Let $\mathbf{1}_\Omega$ be the corresponding indicator function. Then $\mathrm{WF}(\mathbf{1}_\Omega) = \{\alpha d\rho|_\Omega : \alpha \in \mathbb{R} \setminus 0\}$.

Proof. Since $\operatorname{singsupp} \mathbf{1}_\Omega = \partial\Omega$, we only need to consider $\mathrm{WF}(u)$ on the boundary of Ω. Fix $x \in \Omega$, and extend $y_1 = \rho$ to a local coordinate chart $\psi = (y_1, \dots, y_n)$ on a neighborhood $U \subseteq \mathbb{R}^n$ of x, in which $x = (0, \dots, 0)$. In y coordinates, $\mathbf{1}_\Omega$ is locally the indicator of the lower half-plane, $\mathbf{1}_{y_1 \leq 0}$, hence its wavefront set agrees with $\mathrm{WF}(\mathbf{1}_{y_1 \leq 0})$ near 0. Directly computing the Fourier transform of $\mathbf{1}_{y_1 \leq 0}$ gives $-\frac{1}{i\eta_1}\delta_{(\eta_2, \dots, \eta_n)}$, which certainly decays superalgebraically if and only if $\eta_1 \neq 0$. Hence $\mathrm{WF}(\mathbf{1}_{y_1 \leq 0})$ is nonempty and contained in $Z = \{(y, \eta) \mid y_1 =$

$0, \eta_2 = \cdots = \eta_n = 0\}$. But, because of translational symmetry in the $y_1 = 0$ hyperplane, $\mathrm{WF}(\mathbf{1}_{y_1 \leq 0})$ must be exactly Z.

Now, we will peek ahead and use results from section 1.3.2, where it is shown how the wavefront set is modified when we change coordinates from y back to x. Pulling back by ψ can be seen as an Fourier integral operator

$$\psi^* g(x) = \iint e^{i(y-\psi(x))\cdot\eta} g(y)\, dy\, d\eta \tag{1.1.1}$$

with canonical relation $(\psi^{-1}(y), (d\psi)^\top \eta; y, \eta)$. Applying this relation to $\mathrm{WF}(\mathbf{1}_{y_1 \leq 0})$, we obtain $\mathrm{WF}(\psi^{-1}\mathbf{1}_{y_1 \leq 0}) = \{x, cd\rho|_x \mid c \in \mathbb{R},\ x \in \partial\Omega\}$. Since $\psi^{-1}\mathbf{1}_{y_1 \leq 0} = \mathbf{1}_\Omega$ locally, the proof is complete. □

Since microlocal analysis studies the singularities of distributions, two distributions will usually be considered equivalent if their difference is a C^∞ function, while two operators will be considered equivalent if their difference is a *smoothing operator*, meaning that its image is in C^∞.

Given a distribution u, there are a number of operations we can apply: e.g. coordinate changes, pullbacks, pushforwards, and we would like to know how the distribution's wavefront set is modified.

1.2 Pseudodifferential Operators

A very common operator class in analysis of ODEs/PDEs are the pseudodifferential operators, which generalize the class of usual differential operators, and include parametrices for elliptic differential operators and fractional derivative operators, among others.

Definition 1.1 (Pseudodifferential operators on Euclidean space)**.** Let $X \subseteq \mathbb{R}^n$. A pseudodifferential operator P on X is an operator of the form

$$(Pu)(x) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{x\cdot\xi} p(x,\xi)\hat{u}(\xi)\, d\xi, \tag{1.2.1}$$

where $p(x,\xi)$ is a *symbol* satisfying the conditions

$$p \in C^\infty(X \times (\mathbb{R}^n \setminus 0)), \tag{1.2.2}$$

$$\left|\partial_x^\alpha \partial_\xi^\beta p\right| = O\left(|\xi|^{m-|\beta|}\right), \qquad \text{locally uniformly in } X \text{ for some } m \in \mathbb{R} \tag{1.2.3}$$

Here α and β are multi-indices. The second condition is a shorthand way of requiring that for every α, β and every $K \subset\subset X$ there exist constants $C_{\alpha,\beta,K}$ such that $\sup_{x\in K}|\partial_x^\alpha \partial_\xi^\beta p| \leq C_{\alpha,\beta_K}(1+|\xi|)^{m-|\beta|}$.

The growth rate in ξ, m, is the *order* of P, and represents roughly "how many derivatives" are taken by P. When P is an ordinary differential operator, m is equal to the usual order of the operator.

Definition 1.2. We denote the vector space of all pseudodifferential operators of order m on X by $\Psi^m(X)$ and denote by S^m the class of symbols of order m. Also, define $S^{-\infty} = \bigcap_{m\in\mathbb{R}} S^{-m}$.

Note. An operator whose symbol p belongs to $S^{-\infty}$ is smoothing (by an integration by parts argument which will appear later in the proof of Proposition 1.2), so typically we consider any symbols equivalent if they are equal modulo $S^{-\infty}$.

1.2.1 Symbol Classes

The symbol class p given here is the most common symbol class, but others can be considered, where the degree of decay in ξ depends on the order of the derivatives in x and ξ. If $|\partial_x^\alpha \partial_\xi^\beta p| = O\left(|\xi|^{m+\delta|\alpha|-\rho|\beta|}\right)$ locally uniformly in X, then we call p a symbol of type (ρ,δ).

As convenient notation, the operator corresponding to symbol $p(x,\xi)$ is often denoted $p(x,D)$ in view of the fact that the symbol ξ_j corresponds to the pseudodifferential operator $D_j = -i\partial_j$. In general, if $p(x,\xi)$ has a separated representation $p(x,\xi) = \sum_k \alpha_k(x)\beta_k(\xi)$ then $p(x,D) = \sum_k \alpha_k(x)\beta_k(D)$ where $\beta_k(D)$ is defined in the typical functional calculus way via the Fourier transform. However, since the operators x and D do not commute, there

is a not a single canonical definition of $p(x, D)$; the map from the symbol to the operator defined by (1.2.1) is known as *left* or *Kohn-Nirenberg quantization* of p. Other options are *right quantization:*

Definition 1.3 (Right Quantization)**.**

$$(P_R u)(x) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{(x-y)\cdot\xi} p(y, \xi) u(y)\, dy\, d\xi \tag{1.2.4}$$

In right quantization $p(x, D) = \sum_k \alpha_k(x)\beta_k(D)$ if $p(x, \xi) = \sum_k \alpha_k(x)\beta_k(\xi)$; this is adjoint to left quantization. We also have *Weyl quantization*

Definition 1.4 (Weyl Quantization)**.**

$$(P_W u)(x) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{(x-y)\cdot\xi} p\left(\frac{x+y}{2}, \xi\right) u(y)\, dy\, d\xi \tag{1.2.5}$$

These quantizations can be found in [27]. We could envelope all these cases into one by allowing general *amplitude* $a(x, \xi, y)$ in place of the symbol, where a satisfies the symbol decay of (1.2.3), but where derivatives in both x and y are taken (not affecting the exponent of $|\xi|$ in the big-O notation). It turns out that these general amplitudes actually afford no extra generality from a microlocal point of view, since for any amplitude there exists a corresponding symbol p such that $a(x, D, y) = p(x, D)$ modulo smoothing operators:

Proposition 1.2. *For an amplitude $a(x, \xi, y)$, define*

$$a(x, D, y) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{(x-y)\cdot\xi} a(x, \xi, y) u(y)\, dy\, d\xi. \tag{1.2.6}$$

Then there is a $p(x, \xi) \in S^m$ (unique modulo $S^{-\infty}$) such that $a(x, D, y) - p(x, D)$ is smoothing. It can be defined by its asymptotic expansion

$$p(x, \xi) \sim \sum_\alpha \frac{1}{i^{|\alpha|} \cdot \alpha!} \partial_y^\alpha \partial_\xi^\alpha a(x, \xi, y)\Big|_{y=x}. \tag{1.2.7}$$

To make sense of the last part, we need to define what an asymptotic expansions for symbols is, as used in (1.2.7). Before changing gears, we note that this proposition reflects a general theme of Fourier integral operators, of which pseudodifferential operators are a specialization: the amplitude can be thought of in some sense as being defined on the Lagrangian associated with the operator, which for pseudodifferential operators is the diagonal of $(T^*X\backslash 0)\times(T^*X\backslash 0)$, diffeomorphic to $T^*X\backslash 0$. Modifying $a(x,\xi,y)$ away from the diagonal $x=y$ perturbs $a(x,D,y)$ by a smoothing operator, via a stationary phase argument.

1.2.2 Asymptotic Expansions

Oftentimes when dealing with pseudodifferential operators or Fourier integral operators it is useful to be able to separate the symbol into homogeneous terms of different orders. Often also, these forms arise from methods that do not guarantee convergent sums, but only asymptotic expansions:

Definition 1.5 (Asymptotic Expansions)**.** Let $p\in S^{m_0}$, and suppose $p_{m_j}\in S^{m_j}$, $m_0>m_1>m_2>\cdots$. Then

$$p\sim\sum_{j=0}^{\infty}p_{m_j} \tag{1.2.8}$$

if and only if $p-\sum_{j=0}^{J}p_{m_j}\in S^{m_{J+1}}$ for all $j\in\mathbb{N}$.

Despite the lack of convergence, any given asymptotic expansion actually defines a unique symbol $p\in S^{m_0}$ (modulo $S^{-\infty}$).

Proposition 1.3 (Existence and Uniqueness for Asymptotic Expansions)**.** *Let $p_{m_j}\in S^{m_j}$, $m_0>m_1>m_2>\cdots$ be given. Then there exists a symbol $p\in S^m$ such that $p\sim\sum_{j=0}^{\infty}p_{m_j}$, uniquely defined modulo $S^{-\infty}$.*

Proof. Uniqueness is easy, since if p,q are two symbols (of any order) both asymptotic to $\sum_{j=0}^{\infty}p_{m_j}$, then for each J, $p-q=\sum_{j=0}^{J}p_{m_j}-\sum_{j=0}^{J}p_{m_j}=0$, modulo $S^{m_{J+1}}$. Since this is true for all J, $p=q$ modulo $S^{-\infty}$.

If $\sum p_{m_j}$ were a convergent series in $C^\infty(X \times (\mathbb{R}^n \setminus \{0\}))$, we could simply define p to be its sum. If not, it can be shown that we can cut off the p_{m_j} away from the origin, which perturbs the p_{m_j} by a symbol in $S^{-\infty}$ to make the series summable. Having performed the cutoffs, define $p = \sum_{j=0}^{\infty} p'_{m_j}$. □

There is no homogeneity requirement for the p_{m_j} in an asymptotic expansion. Adding this in leads to the polyhomogeneous symbols, which are those having asymptotic expansions in homogeneous symbols with orders descending by 1:

Definition 1.6 (Polyhomogeneous Symbol)**.** A homogeneous symbol p is one of the form

$$p(x,\xi) \sim \sum_{j=0}^{\infty} p_{m-j}(x,\xi) \tag{1.2.9}$$

$$p_{m-j}(x,\xi) \in S^{m-j}, \quad \text{positively homogeneous of degree } m-j \text{ in } \xi \tag{1.2.10}$$

The class of polyhomogeneous symbols is denoted S^m_{phg} in Hörmander and is invariant under addition, composition, and transposition. Returning to the general amplitudes discussed earlier, for a homogeneous amplitude $a(x,y,\xi)$, the corresponding symbol is polyhomogeneous.

With asymptotic expansions defined, we can prove Proposition 1.2. One corollary of this proposition is that if $a(x,y,\xi)$ is homogeneous in ξ, the corresponding symbol is polyhomogeneous.

Proof of Proposition 1.2. As mentioned earlier, the idea is that only the values of $a(x,\xi,y)$ and its derivatives along the diagonal $x = y$ (the critical set for this oscillatory integral) contribute to the operator, modulo smoothing operators. Let k be a positive integer. By Taylor's theorem with remainder, applied in the y variables and centered at $y = x$,

$$a(x,\xi,y) = \sum_{|\alpha|\le k} a_\alpha(x,\xi,y) + \sum_{|\beta|=k+1} (y-x)^\beta R_\beta, \tag{1.2.11}$$

where

$$a_\alpha(x,\xi,y) = \frac{1}{\alpha!}(y-x)^\alpha \partial_y^\alpha a(x,\xi,y)\Big|_{y=x}, \tag{1.2.12}$$

$$R_\beta(x,\xi,y) = \frac{|\beta|}{\beta!}\int_0^1 (1-t)^k (\partial_y^\beta a)(x,\xi,x+t(y-x))\,dt. \tag{1.2.13}$$

The a_α parts we will transform into the symbol, while the R_β parts become the lowest-order terms. Note that each a_α satisfies the symbol decay conditions. First, we pin down the contribution of the a_α terms to the operator.

$$a_\alpha(x,D,y) = \frac{1}{(2\pi)^n}\iint \frac{1}{\alpha!} e^{i(x-y)\cdot\xi}(y-x)^\alpha \partial_y^\alpha a(x,\xi,y)\Big|_{y=x} f(x)\,dx\,d\xi \tag{1.2.14}$$

$$= \frac{1}{(2\pi)^n}\iint \frac{1}{\alpha!} i^{|\alpha|}\partial_\xi^\alpha e^{i(x-y)\cdot\xi}\partial_y^\alpha a(x,\xi,y)\Big|_{y=x} f(x)\,dx\,d\xi, \tag{1.2.15}$$

and applying integration by parts,

$$= \frac{1}{(2\pi)^n}\iint e^{i(x-y)\cdot\xi}\frac{1}{i^{|\alpha|}\cdot\alpha!}\partial_\xi^\alpha\partial_y^\alpha a(x,\xi,y)\Big|_{y=x} f(x)\,dx\,d\xi. \tag{1.2.16}$$

The right-hand side is now a standard left-quantization-defined pseudodifferential operator. Likewise,

$$\left((y-x)^\beta R_\beta\right)(x,D,y) = \frac{(-1)^{|\beta|}|\beta|}{(2\pi)^n\beta!}\iint e^{i(x-y)\cdot\xi}\partial_\xi^\beta R_\beta(x,y,\xi)\,dx\,d\xi. \tag{1.2.17}$$

Now analyzing the R_β in (1.2.13) shows that they are also amplitudes of the same order as m. Therefore each $\left((y-x)^\beta R_\beta\right)(x,D,y)$ is an amplitude of order $m-|\beta|$ (and thus so is the second sum in (1.2.11)).

Now, we compare $a(x,D,y)$ with the operator whose symbol $q(x,\xi)$ is defined by asymptotic expansion

$$q(x,\xi) \sim \sum_\alpha \frac{1}{i^{|\alpha|}\cdot\alpha!}\partial_y^\alpha\partial_\xi^\alpha a(x,\xi,y)\Big|_{y=x}. \tag{1.2.18}$$

By (1.2.16), $a(x,D,y)-q(x,D)$ is a pseudodifferential operator defined by an amplitude of order $m-|\beta|$ for all β. This implies it is a smoothing operator: after taking k derivatives of

$$Af(x) = \int e^{i(x-y)\cdot\xi}a(x,y,\xi)f(y)\,dy\,d\xi \tag{1.2.19}$$

we increase the growth order of the integrand in θ by k, but since $a(x, y, \xi)$ decays faster than any power of ξ, the integral is still convergent. This finishes the proof.

□

1.2.3 Symbol Calculus

One of the most convenient properties of pseudodifferential operators is their symbol calculus, which allows us to compute symbols of the composition of two pseudodifferential operators, the (microlocal) inverse of an operator, and more.

The transpose of a pseudodifferential operator is easily expressed as a general amplitude, which leads us to a formula for its symbol via Proposition 1.2.

Proposition 1.4. *For any $p(x, \xi) \in S^m$, the transpose of $P = p(x, D)$ is a pseudodifferential operator defined by amplitude $a(x, \xi, y) = p(y, -\xi)$. Modulo smoothing operators, its (Kohn-Nirenberg) symbol has asymptotic expansion*

$$\sigma_{P^\top} \sim \sum_\alpha \frac{i^{|\alpha|}}{\alpha!}(\partial_x^\alpha \partial_\xi^\alpha p)(x, -\xi). \tag{1.2.20}$$

Similarly, the adjoint P^* is defined by amplitude $a(x, \xi, y) = \overline{p(y, \xi)}$. Its symbol has the asymptotic expansion

$$\sigma_{P^*} \sim \sum_\alpha \frac{(-i)^{|\alpha|}}{\alpha!}\overline{(\partial_x^\alpha \partial_\xi^\alpha p)(x, \xi)}. \tag{1.2.21}$$

Proof. The transpose of P's kernel is formally

$$\frac{1}{(2\pi)^n}\int e^{i(y-x)\cdot\xi} p(y, \xi)\, d\xi. \tag{1.2.22}$$

Substituting $\xi = -\xi$ yields a pseudodifferential operator defined by amplitude $p(y, -\xi)$, and applying Proposition 1.2 yields (1.2.20). Similarly, P^*'s kernel is obtained by taking the conjugate transpose of P's kernel, giving

$$\frac{1}{(2\pi)^n}\int e^{-i(y-x)\cdot\xi}\overline{p(y, \xi)}\, d\xi. \tag{1.2.23}$$

This yields amplitude $\overline{p(y, \xi)}$, and Proposition 1.2 implies (1.2.21). □

Now that the transpose is defined, it is not hard to find the symbol of the composition of two pseudodifferential operators, by expressing one in transpose form. The composition is as elegant as can be hoped for:

Proposition 1.5. *Let $P = p(x, D) \in \Psi^{m_1}(X)$, $Q = q(x, D) \in \Psi^{m_2}(X)$ be two pseudodifferential operators on X. Then PQ is a pseudodifferential operator of order $m_1 + m_2$. Its symbol has the asymptotic expansion*

$$\sigma_{PQ} \sim \sum_{\alpha} \frac{1}{\alpha!} (\partial_\xi^\alpha p)(\partial_x^\alpha q). \tag{1.2.24}$$

The lower-order terms of the asymptotic expansion reflect the non-commutativity of x and D.

Proof. Write $q^\top$ for the symbol of $Q^\top$. Then

$$Pu(x) = \frac{1}{(2\pi)^n} \int e^{ix\cdot\xi} p(x, \xi) \hat{u}(\xi)\, d\xi, \tag{1.2.25}$$

$$Q^\top u(z) = \frac{1}{(2\pi)^n} \iint e^{i(z-y)\cdot\xi} q^\top(z, \xi) u(y)\, dy\, d\xi, \tag{1.2.26}$$

$$Qu(y) = \frac{1}{(2\pi)^n} \iint e^{i(y-z)\cdot\xi} q^\top(z, -\xi) u(z)\, dz\, d\xi, \tag{1.2.27}$$

$$\widehat{Qu}(\xi) = \int e^{-iz\cdot\xi} q^\top(z, -\xi) u(z)\, dz. \tag{1.2.28}$$

Composing P and Q from the first and last equations gives

$$(PQu)(x) = \frac{1}{(2\pi)^n} \int e^{i(x-z)\cdot\xi} p(x, \xi) q^\top(z, -\xi) u(z)\, dz\, d\xi. \tag{1.2.29}$$

The product of two symbols is again a symbol (simply by applying the product rule), so (1.2.29) demonstrates that the PQ is a pseudodifferential operator, and its amplitude has order $m_1 + m_2$. Now we can apply Proposition 1.2, gaining

$$\sigma_{PQ} \sim \sum_{\alpha} \frac{(-i)^{|\alpha|}}{\alpha!} \partial_\xi^\alpha \partial_z^\alpha (p(x, \xi) q^\top(z, -\xi)) \Big|_{z=x}. \tag{1.2.30}$$

Combining this with the asymptotic expansion for $q^\top$, $\sum_\beta \frac{i^{|\beta|}}{\beta!} (\partial_z^\beta \partial_\xi^\beta q)(z, -\xi)$, we get

$$\sigma_{PQ} \sim \sum_{\alpha,\beta} \frac{(-i)^{|\alpha|-|\beta|}}{\alpha!\beta!} \partial_\xi^\alpha \partial_z^\alpha \left(p(x, \xi) (\partial_z^\beta \partial_\xi^\beta q)(z, \xi) \right) \Big|_{z=x}. \tag{1.2.31}$$

Expanding the ∂_ξ^α derivative with the product rule, we apply $\gamma \leq \alpha$ of the derivatives on p; the remaining on the derivative of q:

$$\sigma_{PQ} \sim \sum_{\alpha,\beta} \sum_{\gamma \leq \alpha} \frac{(-i)^{|\alpha|-|\beta|}}{\alpha!\beta!} \binom{\alpha}{\gamma} (\partial_\xi^\gamma p)(\partial_x^{\alpha+\beta} \partial_\xi^{\alpha+\beta-\gamma} q). \tag{1.2.32}$$

Replacing α by $\alpha + \gamma$, then by $\delta = \alpha + \beta$,

$$\sigma_{PQ} \sim \sum_{\alpha,\beta,\gamma} \frac{(-i)^{|\alpha|-|\beta|+|\gamma|}}{\alpha!\beta!\gamma!} (\partial_\xi^\gamma p)(\partial_x^{\alpha+\beta+\gamma} \partial_\xi^{\alpha+\beta} q) \tag{1.2.33}$$

$$\sim \sum_{\delta,\gamma} \sum_{\beta \leq \delta} \frac{(-i)^{|\delta|+|\gamma|}}{\gamma!\delta!} (-1)^{|\beta|} \binom{\delta}{\beta} (\partial_\xi^\gamma p)(\partial_x^{\delta+\gamma} \partial_\xi^\delta q). \tag{1.2.34}$$

Unless $|\delta| = 0$, the alternating sum $\sum_{\beta \leq \delta} (-1)^{|\beta|} \binom{\delta}{\beta}$ is zero. For, choosing some i such that $\beta_i \neq 0$, and letting ϵ be the multi-index with $\epsilon_i = 1$, $\epsilon_j = 0$ for $j \neq i$ we have $\binom{\delta}{\beta} = \binom{\delta-\epsilon}{\beta-\epsilon} + \binom{\delta-\epsilon}{\beta}$, so each nonzero $\binom{\delta-\epsilon}{\zeta}$ appears twice in the sum, once with positive sign and once with negative sign. Hence

$$\sigma_{PQ} \sim \sum_{\gamma} \frac{(-i)^{|\gamma|}}{\gamma!} (\partial_\xi^\gamma p)(\partial_x^\gamma q). \tag{1.2.35}$$

□

One of the major uses for pseudodifferential operators is providing approximate inverses for elliptic differential operators. A pseudodifferential operator $p(x, D)$ is said to be elliptic if its symbol p is suitably bounded away from zero.

Definition 1.7. $p(x, D)$ is *elliptic* if $p(x, \xi) = \omega(|\xi|^m)$ locally uniformly in x, where m is the order of p. In other words, for every $K \subset\subset X$ there exist constants $a, c > 0$ such that

$$\sup_{x \in K} |p(x, \xi)| \geq c\,|\xi|^{\,m} \quad \text{if } |\xi| \geq a. \tag{1.2.36}$$

We can now prove:

Proposition 1.6. *For every elliptic $P \in \Psi^m(X)$, there is a pseudodifferential parametrix $Q \in \Psi^{-m}(X)$ satisfying $PQ = \mathrm{Id}$ modulo a smoothing operator.*

Proof. Let $p(x,\xi)$ be the symbol of P. Let $\mathcal{K} = \{K_1 \subset K_2 \subset \cdots\}$ be an exhaustion of X by compact subsets. For each $K \in \mathcal{K}$, let $a_K, c_K > 0$ be constants satisfying (1.2.36). Let $V_K = \{(x,\xi) : x \in K,\ |\xi| \geq a_K\}$.

Because p is elliptic, $\tilde{q} = \frac{1}{p}$ satisfies the symbol decay conditions on each V_K: we can check that $\partial_x^\alpha \partial_\xi^\beta \tilde{q} = s_{\alpha,\beta} p^{-|\alpha|-|\beta|}$, where $s_{\alpha,\beta}$ is a symbol of order $m\alpha + (m-1)\beta$. Choose a smooth cutoff $\chi(x,\xi)$ with $0 \leq \chi \leq 1$, such that χ is supported on $\bigcup_{K\in\mathcal{K}} V_K$ and the support of $1-\chi$ in ξ is bounded on each K. Then define $q_0 = \chi/p$. Since $q_0 = \tilde{q}$ for sufficiently large ξ, the symbol decay conditions are satisfied, while the cutoff χ ensures $q_0 \in C^\infty$. Hence $q_0 \in S^{-m}$. Let $Q_0 = q_0(x,D)$.

Applying the composition formula, we find that while the symbol of PQ_0 is not 1, its principal symbol is, and so $PQ_0 = \mathrm{Id} - P_1$ for some pseudodifferential operator $P_1 = p_1(x,D)$ of order -1. Having done so, we can repeat the argument with $q_1 = \chi p_1/p$ to find $Q_1 = q_1(x,D)$ such that $PQ_1 = P_1 - P_2$ for some pseudodifferential operator $P_2 = p_2(x,D)$ of order -2, and so on. Continuing similarly, we generate $P_N \in \Psi^{-N}$, $Q_N \in \Psi^{-m-N}$ such that $P(Q_0 + Q_1 + Q_N) = \mathrm{Id} - P_{N+1}$ for each N. While the sum $Q_0 + Q_1 + \cdots$ may not converge, there exists a $Q \in \Psi^{-m}$ whose symbol is asymptotic to $\sum_{j=0}^\infty q_j$. Then for each N,

$$PQ - \mathrm{Id} = (PQ - P(Q_0 + \cdots + Q_N)) + (P(Q_0 + \cdots + Q_N) - \mathrm{Id}) \in \Psi^{-N-1}. \tag{1.2.37}$$

Hence $PQ = \mathrm{Id}$ plus a smoothing operator. □

1.3 Fourier Integral Operators

In microlocal analysis and inverse problems, many operators of study belong to the larger class of Fourier integral operators, which are a wide generalization of pseudodifferential operators. In essence, a Fourier integral operator, or FIO, is an integral operator with kernel (locally) defined by an oscillatory integral. One of their most important and useful features is that they move singularities (the wavefront set) in a prescribed manner. Fourier integral operators include integral transforms such as the Radon transform and its generalizations, and parametrices for hyperbolic equations, as well as pushforwards and pullbacks. As a

result, they have found wide application to inverse problems in imaging. There is a well-established calculus for FIOs, describing the circumstances under which the composition of FIOs is itself an FIO. Like pseudodifferential operators, Fourier integral operators can be described up to leading order by an appropriate principal symbol, although there is not (to the author's knowledge) an analogue to the full symbol possessed by pseudodifferential operators.

1.3.1 Definition and Examples

The kernels of Fourier integral operators are known sometimes as Fourier integral distributions.

Definition 1.8. Let $X \subseteq \mathbb{R}^n$ be open. A local Fourier integral distribution $u \in \mathcal{D}'(X)$ is a distribution defined by

$$\langle u, f\rangle = (2\pi)^{-(n+2N)/4} \int_{\mathbb{R}^N} e^{i\phi(x,\theta)} a(x,\theta) f(x)\, dx\, d\theta, \qquad f \in C_0^\infty(X). \tag{1.3.1}$$

where $a \in C^\infty(X \times \mathbb{R}^N)$ and $\phi \in C^\infty(X \times (\mathbb{R}^N \setminus 0))$ are known as the amplitude and phase functions, respectively. The amplitude a must satisfy the symbol decay conditions

$$\partial_x^\alpha \partial_\theta^\beta a = O(|\theta|^{\mu-\beta}). \tag{1.3.2}$$

locally uniformly on X for some order $\mu \in \mathbb{R}$. The phase ϕ is required to be (positively) homogeneous of degree 1, real-valued (although complex-phase FIOs can also be considered), with $d\phi \neq 0$. Typically, ϕ is also required to be *nondegenerate*, meaning that $d_{(x,\theta)} d_\theta \phi$ has full rank (this implies that the critical manifold, defined later, is a embedded submanifold).

When the order μ of the amplitude is less than $-n$, the integrand of (1.3.1) is absolutely convergent (regardless of the symbol bounds given). To define (1.3.1) for $\mu \geq -n$, we must use integration by parts and the symbol decay properties of a, as well as the condition $d\phi \neq 0$. Let L be the first-order operator on $T^*X \setminus 0$:

$$L = -i \frac{\frac{\partial\phi}{\partial x}\frac{\partial}{\partial x} + |\theta|^2 \frac{\partial\phi}{\partial\theta}\frac{\partial}{\partial\theta}}{\left|\left(\frac{\partial\phi}{\partial x}, |\theta|\frac{\partial\phi}{\partial\theta}\right)\right|^2}. \tag{1.3.3}$$

Since $d\phi \neq 0$, L has finite coefficients, and by construction $L(i\phi) = 1$, so $L(e^{i\phi}) = e^{i\phi}$. Meanwhile, the coefficient of $\frac{\partial}{\partial x}$ is homogeneous of order -1 in θ and the coefficient of $\frac{\partial}{\partial \theta}$ is order 0, so the symbol conditions imply that $L^k a$ is $O(\theta^{\mu-k})$ locally uniformly in x. More to the point here, the coefficients of $L^\top$ are homogeneous of the same orders, although $L^\top$ also has a constant term that is homogeneous of order -1 in θ, so $(L^\top)^k a$ is $O(\theta^{\mu-k})$ locally uniformly in x. Formally integrating (1.3.1) by parts

$$\langle u, f\rangle = (2\pi)^{-(n+2N)/4} \int e^{i\phi(x,\theta)} a(x,\theta) f(x)\, dx\, d\theta \tag{1.3.4}$$

$$= (2\pi)^{-(n+2N)/4} \int L^k e^{i\phi(x,\theta)} a(x,\theta) f(x)\, dx\, d\theta \tag{1.3.5}$$

$$= (2\pi)^{-(n+2N)/4} \int e^{i\phi(x,\theta)} (L^\top)^k a(x,\theta) f(x)\, dx\, d\theta. \tag{1.3.6}$$

By choosing $k > \mu + n$, the final integral (1.3.6) is absolutely convergent. By integration by parts, it is also independent of k. We take this convergent integral as the definition of the distribution u given by the oscillatory integral (1.3.1).

Definition 1.9. A local Fourier integral operator $A\colon C_0^\infty(Y) \to \mathcal{D}'(X)$ is a integral operator whose Schwartz kernel is a Fourier integral distribution $\int e^{i\phi(x,y,\theta)} a(x,y,\theta)\, d\theta$ on $X \times Y$.

The *order* of A is defined to be $\frac{1}{4}(-\dim X - \dim Y + 2\dim \theta)$, and measures "how many derivatives A takes"; this can be made precise, at least for FIOs associated with canonical graphs, as will be discussed later.

Note. A priori, A maps $C_0^\infty(Y)$ to $\mathcal{D}'(X)$, since its Schwartz kernel is a distribution; however, it turns out that after studying the effect of Fourier integral operators on the wavefront sets of distributions more carefully, the domain of A can be extended to distributions with wavefront sets outside a certain "bad set" which can be computed from ϕ.

Example. Any pseudodifferential operator $p(x, D)$ is a Fourier integral operator whose order is equal to the degree of the symbol p. The phase is $\phi(x, y, \xi) = (x - y) \cdot \xi$, which satisfies the conditions for a nondegenerate phase function.

Example. Pullbacks and pushforwards of distributions by smooth maps $\psi\colon Y \to X$ are Fourier integral operators:

$$\psi^* f(y) = f(\psi(y)) = \int e^{i(\psi(y)-x)\cdot\xi} f(x)\, dx\, d\xi; \tag{1.3.7}$$

$$\psi_* g(x) = \int e^{i(\psi(y)-x)\cdot\xi} g(y)\, dy\, d\xi. \tag{1.3.8}$$

The phases are identical in both cases and nondegenerate.

Example. The solution operator for a constant-speed wave equation is a sum of two FIOs. For simplicity taking wave speed 1, the solution to $(\partial_t^2 - \Delta_x^2)u = 0$ with initial data $(e^{i\langle\cdot,\xi\rangle}, 0)$, $f_\xi(x) = e^{ix\cdot\xi}$, is the sum $u(t,x) = \frac{1}{2}[e^{i(x\cdot\xi+t|\xi|)} + e^{i(x\cdot\xi-t|\xi|)}]$. For Cauchy data $(f, 0)$ with general f, we take the Fourier transform and use linearity to get

$$u(x) = \frac{1}{2}\int e^{i[x\cdot\xi+t|\xi|]}\hat{f}(\xi)\, d\xi + \frac{1}{2}\int e^{i[x\cdot\xi-t|\xi|]}\hat{f}(\xi)\, d\xi \tag{1.3.9}$$

$$= \frac{1}{2}\int e^{i[x\cdot\xi+t|\xi|-y\cdot\xi]} f(y)\, dy\, d\xi + \frac{1}{2}\int e^{i[x\cdot\xi-t|\xi|-y\cdot\xi]} f(y)\, dy\, d\xi \tag{1.3.10}$$

These two integrals are FIOs, as usual with nondegenerate phases (and amplitude 1).

Example. The Radon transform $R\colon C_0^\infty(\mathbb{R}^n) \to C^\infty(\mathbb{R}\times\mathbb{S}^{n-1})$ integrates functions over hyperplanes in Euclidean space:

$$Rf(s,\omega) = \int_{x\cdot\omega=s} f(x)\, dS. \tag{1.3.11}$$

The Fourier slice theorem implies that $\mathcal{F}_s R(\sigma,\omega) = \hat{f}(\sigma\omega)$, where $\mathcal{F}_s$ is the Fourier transform of Rf in s. By the Fourier inversion formula,

$$Rf(s,\omega) = \frac{1}{2\pi}\int e^{is\sigma}\hat{f}(\sigma\omega)\, d\sigma \tag{1.3.12}$$

$$= \frac{1}{2\pi}\int e^{i[s\sigma-x\cdot\sigma\omega]} f(x)\, dx\, d\sigma, \tag{1.3.13}$$

which gives us an oscillatory integral with phase $\phi(x,s,\omega) = \sigma(s - x\cdot\omega)$ and amplitude $1/2\pi$. This is essentially a Dirac delta distribution on the manifold $\{s = x\cdot\omega\}$. Since the codomain $Y = \mathbb{R}\times\mathbb{S}^{n-1}$ is a manifold, this does not exactly fit the local definition of a FIO given earlier, but we will shortly define FIOs on manifolds.

We can also generalize the Radon transform, integrating over families of hypersurfaces of a space X. If the hypersurfaces are level sets of a critical-point free smooth function $h\colon X \times \Omega \to \mathbb{R}$, then the associated Radon transform is

$$R_h f(s,\omega) = \int_{h(x,\omega)=s} f(x)\, dS \tag{1.3.14}$$

Fixing ω, choose local coordinates $\tilde{x}$ in which $h(\cdot,\omega) = \tilde{x}\cdot s\omega$; then $Rf(s,\omega)$ is the standard Radon transform so we find that $Rf(s,\omega) = \int e^{i\sigma(s-\tilde{x}\cdot\omega)} f(\tilde{x})\, d\tilde{x}\, d\omega$ for f supported in the coordinate patch. Changing coordinates back to x gives us

$$Rf(s,\omega) = \int e^{i\sigma(s-h(x,\omega))} f(x) \det \frac{\partial \tilde{x}}{\partial x}\, dx\, d\omega. \tag{1.3.15}$$

We can check that $\det \frac{\partial \tilde{x}}{\partial x} = 1$. This is true for all ω, so again we find an FIO.

1.3.2 FIOs and Wavefront Sets

One of the most salient and useful properties of FIOs is that they move singularities in a prescribed way. To find out how $\mathrm{WF}(f)$ and $\mathrm{WF}(Af)$ are related for an FIO A, we need to look at the wavefront set of the kernel of A. It turns out to be contained in a certain Lagrangian submanifold $\Lambda_\phi \in T^*X$ which is related to the critical points of the phase ϕ with respect to the frequency variables θ. The wavefront set of the kernel comes from the method of stationary phase, which will tell us that $\mathcal{F}_{(x,y)} \int e^{i\phi(x,y,\theta)} a(x,y,\theta)\, d\theta$ decays rapidly in $|\theta|$ away from the critical points of its phase. Most of this section follows Duistermaat [18].

Before embarking on finding the wavefront set of A's kernel, we find the accompanying manifolds C_ϕ and Λ_ϕ. These are defined for any Fourier integral distribution:

Definition 1.10. Let $\phi(x,\theta)$ be a phase function, u a Fourier integral distribution. The *critical manifold* of ϕ is

$$C_\phi = \{(x,\theta) \mid d_\theta \phi(x,\theta) = 0\} \subseteq X \times \mathbb{R}^N. \tag{1.3.16}$$

The *Lagrangian manifold* is

$$\Lambda_\phi = \left\{(x,\xi) \in T^*X \;\middle|\; \exists \theta \in \mathbb{R}^N \colon (x,\theta) \in C_\phi,\ \xi = d_x\phi(x,\theta)\right\}. \tag{1.3.17}$$

Λ_ϕ is said to be parameterized by ϕ. When u is the kernel of a Fourier integral operator A, Λ_ϕ is said to be the *canonical relation* of A.

The nondegeneracy condition imposed earlier on the phase ($d_{(x,\theta)}d_\theta\phi \neq 0$) implies that C_ϕ is an embedded submanifold of $X \times \mathbb{R}^N$. Λ_ϕ is the image of C_ϕ under the map $T\colon C_\phi \to T^*X$; $T(x,\theta) \mapsto (x, d_x\theta)$. A quick argument (see Duistermaat [1]) shows that dT is injective, so Λ_ϕ is an immersed submanifold in T^*X. The dimensions of C_ϕ and Λ_ϕ are both n. As its name implies, Λ is a Lagrangian submanifold of T^*X with respect to the canonical symplectic form on T^*X, $\omega = dx^i \wedge d\xi_i$, meaning that $\omega|_\Lambda = 0$ and $\dim \Lambda = n$.

For Fourier integral operators $A\colon C^\infty(Y) \to \mathcal{D}'(X)$, we have $C_\phi \subset X \times Y \times \mathbb{R}^N$ and $\Lambda_\phi \subseteq T^*X \times T^*Y$.

Proposition 1.7. *If $u = \int e^{i\phi} a\, d\theta$ is any Fourier integral distribution,* $\mathrm{WF}(u) \subseteq \Lambda_\phi$.

Proof. Let $(x_0, \xi_0) \in (T^*X \setminus 0) \setminus \Lambda_\phi$. We would like to show that $\widehat{\rho u}(t\xi) = O(t^{-N})$ for all N when ξ in a conic neighborhood of ξ_0, for some smooth cutoff ρ supported near x_0. Note that

$$(\rho u)(x) = \int e^{i\phi(x,\theta)}(\rho a)(x,\theta)\, d\theta \tag{1.3.18}$$

$$\widehat{\rho u}(t\,|\xi|) = \int e^{i\phi(x,\theta) - x\cdot t\xi}(\rho a)(x,\theta)\, d\theta\, dx \tag{1.3.19}$$

and after substituting $\theta = t\theta'$ and using homogeneity,

$$\widehat{\rho u}(t\,|\xi|) = t^N \int e^{it[\phi(x,\theta') - x\cdot\xi]}(\rho a)(x, t\theta')\, d\theta'\, dx. \tag{1.3.20}$$

We need to study the asymptotic behavior of this integral. It falls into the general class of oscillatory integrals $\int e^{itf(z,a)} g(z,a,t)\, dz$, where $z \in Z \subseteq \mathbb{R}^{N_Z}$, and $a \in A \subseteq N_A$ are arbitrary parameters. For f and g, suppose $f \in C^\infty(Z \times A)$ and $g \in C^\infty(Z \times A \times \mathbb{R}_+)$, and $g = 0$ for z outside some $K \subset\subset Z$. Also, we need to assume that $\sup_{z,a} |\partial_z^j g| = O(t^{m+j\delta})$ for some $\delta < -\frac{1}{2}$.

If (for each a), f has a unique nondegenerate critical point with respect to z, say $z_0(a)$, and $g(z_0(a), a, t)$ is nonzero for all a, t, the method of stationary phase says that

$$\int e^{itf(z,a)} g(z,a,t)\, dz = e^{itf(z_0(a),a)} g(z_0(a), a, t) c(a) t^{N_z/2} + o(t^{N_z/2}). \tag{1.3.21}$$

for $c(a)$ a constant depending on the Hessian $d_z^2 f$ at $z_0(a)$ (the critical point being nondegenerate means that the Hessian is nonsingular for each a). But, the important part for our original integral is that we would not get the rapid decay that indicates smoothness in the ξ direction. On the other hand, if f has no critical points, then instead the integral $e^{itf(z,a)}g(z,a,t)\,dz$ decays faster than any negative power of t.

Returning to our problem, $f = \phi(x,\theta') - x\cdot\xi$, and its derivatives are (dropping primes from θ'):

$$d_\theta[\phi(x,\theta) - x\cdot\xi] = d_\theta\phi, \tag{1.3.22}$$

$$d_\xi[\phi(x,\theta) - x\cdot\xi] = d_x\phi - \xi. \tag{1.3.23}$$

So, the critical points of the phase f are exactly (x,θ) for which $d_\theta(x,\theta) = 0$ ($(x,\theta)\in C_\phi$) and $\xi = d_x\phi(x,\theta)$.

Choose neighborhoods $U \subseteq \overline{U} \subseteq U' \subseteq X$ of x_0 and a conic neighborhood $V \subseteq \mathbb{R}^n \setminus \{0\}$ of ξ_0 such that $(U'\times V)\cap\Lambda_\phi = \emptyset$. Then choose $\rho \in C_0^\infty(U)$ such that $\rho(x_0)\neq 0$. Also choose $\tilde\rho \in C_0^\infty(U')$ with $\tilde\rho\big|_U = 1$. For any $\xi\in V$, the phase of $\widehat{\rho u}(t\xi)$ has no critical points on the support of ρa. Hence we expect that (1.3.20) is rapidly decaying for ξ in the neighborhood V.

To prove it [3], we introduce a particular first-order differential operator L (as in the definition of oscillatory integrals) and apply integration by parts. Since df is nonzero on $U'\times\mathbb{R}^N$ for $\xi\in V$, define

$$L = \frac{\tilde\rho(x)}{i}\frac{\widetilde{\nabla} f\cdot\nabla_{x,\theta}}{\widetilde{\nabla} f\cdot\nabla f}, \qquad\qquad \widetilde{\nabla} f = \left(\nabla_x f, |\theta|^2\nabla_\theta f\right). \tag{1.3.24}$$

The cutoff $\tilde\rho$ ensures that L has C^∞ coefficients everywhere; $\widetilde{\nabla} f\cdot\nabla f \neq 0$ on U' since $df\neq 0$. By construction, $Lf = \frac{1}{i}$, so $Le^{itf} = te^{itf}$, and we can try using the same integration by

parts trick as before:

$$\widehat{\rho u}(t\xi) = \int e^{itf(x,\theta,\xi)}(\rho a)(x,t\theta)\,d\theta\,dx \tag{1.3.25}$$

$$= \int t^{-k}L^k e^{itf(x,\theta,\xi)}(\rho a)(x,t\theta)\,d\theta\,dx \tag{1.3.26}$$

$$= t^{-k}\int e^{itf(x,\theta,\xi)}(L^\top)^k\Big[(\rho a)(x,t\theta)\Big]\,d\theta\,dx. \tag{1.3.27}$$

Now we look at $L^\top$, where the reason for using $\widetilde{\nabla} f$ in (1.3.24) appears. The denominator $\nabla f \cdot \widetilde{\nabla} f$ is just $|\nabla_x \phi|^2 + |\theta|^2|\nabla_\theta \phi|^2$, which is homogeneous of order 2 in θ, since $\nabla_x \phi$, $\nabla_\theta \phi$ are homogeneous of orders 1 and 0, respectively. If we write $L = b \cdot \nabla_x + c \cdot \nabla_\theta$ then, b and c are homogeneous of orders -1 and 0, respectively. Taking the transpose,

$$L^\top = -b \cdot \nabla_x - c \cdot \nabla_\theta + (\nabla_x \cdot b + \nabla_\theta \cdot c). \tag{1.3.28}$$

All the coefficients but c are homogeneous of order -1, and since ρa is a symbol, $\nabla_\theta(\rho a)$ is a symbol of degree one lower. Hence if $a \in S^m$ then $(L^\top)^k(\rho a) \in S^{m-k}$. If $m - k < -N$, then the integrand in (1.3.27) is absolutely convergent (since $(L^\top)\rho a = o(|\theta|^{-n})$ and it is compactly supported in x), and bounded for $\xi \in V$, so we have $\widehat{\rho u}(t\xi) = O(t^{-k})$ uniformly on V. This finishes the proof. □

Now that we have the wavefront set of the kernel of A, we need to understand what this tells us about the relationship between $\mathrm{WF}(f)$ and $\mathrm{WF}(Af)$. First, though, we need to extend the domain of A beyond smooth functions. However, in general an FIO cannot be extended to every distribution. Accordingly, we define some appropriate spaces of distributions:

Definition 1.11. Let $X \subseteq \mathbb{R}^{n_X}$ be open, and let Γ be a closed conic subset of $T^*X \setminus 0$. Then

$$\mathcal{D}'_\Gamma(X) = \{u \in \mathcal{D}'(X) \mid \mathrm{WF}\, u \subseteq \Gamma\}. \tag{1.3.29}$$

Next, we introduce a modified version the wavefront set of the kernel, the *wavefront relation*, and we also introduce a few extra sets:

Definition 1.12. Let $A\colon C^\infty(Y) \to \mathcal{D}'(X)$ be a linear operator with Schwartz kernel K_A. Then the wavefront relation is

$$\mathrm{WF}'(K_A) = \{(x,\xi;y,-\eta) \in T^*X \times T^*Y \mid (x,\xi;y,\eta) \in \mathrm{WF}(K_A)\}. \tag{1.3.30}$$

Also, define the following sets:

$$\mathrm{WF}'_{in}(K_A) = \left\{(y,\eta) \in T^*Y \setminus 0 \;\middle|\; \exists x \in X : (x,0,y,\eta) \in \mathrm{WF}'(K_A)\right\}, \tag{1.3.31}$$

$$\mathrm{WF}'_{out}(K_A) = \left\{(x,\xi) \in T^*X \setminus 0 \;\middle|\; \exists y \in Y : (x,\xi,y,0) \in \mathrm{WF}'(K_A)\right\}. \tag{1.3.32}$$

Abusing notation as mathematicians are wont, we write $\mathrm{WF}'(A) = \mathrm{WF}'(K_A)$ and similarly for WF'_{in}, WF'_{out}.

The name "wavefront relation" reflects the fact that $\mathrm{WF}'(A)$ can be thought of as a relation from T^*Y to T^*X. It turns out, as we will see shortly, that this relation relates $\mathrm{WF}(f)$ and $\mathrm{WF}(Af)$.

We have the following key facts (not proven here):

Theorem 1.8. *Let $X \subseteq \mathbb{R}^{n_X}$, $Y \subseteq \mathbb{R}^{n_Y}$. Let $A\colon C^\infty(Y) \to \mathcal{D}'(X)$ be a linear operator. Let Γ be a closed conic set in $T^*X \setminus 0$ disjoint from $\mathrm{WF}_{in}(T)$. Then A extends continuously to a map $A\colon \mathcal{D}'_\Gamma(X) \to \mathcal{D}'(X)$.*

Theorem 1.9. *Let $A\colon C^\infty(Y) \to \mathcal{D}'(X)$ be a linear operator, and let $u \in \mathcal{D}'(X)$ with $\mathrm{WF}(u) \cap \mathrm{WF}'_{in}(A) = \emptyset$. Then*

$$\mathrm{WF}'(Au) \subseteq \Big(\mathrm{WF}'(A) \circ \mathrm{WF}(u)\Big) \cup \mathrm{WF}'_{out}(A). \tag{1.3.33}$$

Looking over Theorem 1.9, we see that $\mathrm{WF}'_{in}(A)$ can be thought of the set of "bad" singularities which A cannot be applied to. On the other end, $\mathrm{WF}'_{out}(A)$ are "extra" singularities which A can create without any singularities present in the input.

Combining Theorem 1.9 with Proposition 1.7, we achieve our goal:

Corollary 1.10. *Let $A\colon C^\infty(Y) \to \mathcal{D}'(X)$ be a Fourier integral operator with phase ϕ, and define Λ'_ϕ, $\Lambda'_{\phi,in}$, $\Lambda'_{\phi,out}$ analogously to (1.12). Then A extends continuously to a map $\mathcal{D}'_\Gamma(X) \to \mathcal{D}'(X)$ for any closed conic $\Gamma \subseteq T^*X \setminus (\Lambda'_{\phi,in} \cup 0)$ and*

$$\mathrm{WF}(Af) \subseteq (\Lambda'_\phi \circ \mathrm{WF}(f)) \cup \Lambda'_{\phi,out}. \tag{1.3.34}$$

1.3.3 Composition

The natural question for a class of operators is whether it is closed under composition. A priori, since an FIO maps C_0^∞ to $\mathcal{D}'$, it is not clear whether we can compose two operators. It turns out that under some geometric conditions on some manifolds associated to the FIOs, this composition is indeed possible.

Theorem 1.11. *Let $X \subseteq \mathbb{R}^{n_X}$, $Y \subseteq \mathbb{R}^{n_Y}$, $Z \subseteq \mathbb{R}^{n_Z}$. Let $A\colon C^\infty(Y) \to \mathcal{D}'(X)$, $B\colon C^\infty(Z) \to \mathcal{D}'(Y)$ be local Fourier integral operators of orders m_1, m_2 respectively.*

$$Af(x) = (2\pi)^{-(n_X+n_Y+2N_1)/4} \int_{\mathbb{R}^{N_1}} \int_Y e^{i\phi_1(x,y,\theta_1)} a(x,y,\theta_1) f(y)\, d\theta_1 \tag{1.3.35}$$

$$Bf(y) = (2\pi)^{-(n_Y+n_Z+2N_2)/4} \int_{\mathbb{R}^{N_2}} \int_Z e^{i\phi_2(y,z,\theta_2)} b(y,z,\theta_2) g(z)\, d\theta_2. \tag{1.3.36}$$

Let Λ_{ϕ_1}, Λ_{ϕ_2} be the associated Lagrangian manifolds. Assume the following conditions hold:

(a) $\Lambda_{\phi_1,in} = \Lambda_{\phi_2,out} = \emptyset$.

(b) $\Lambda_{\phi_1,out} \cap \Lambda_{\phi_2,in} = \emptyset$.

(c) $\Lambda_{\phi_1} \times \Lambda_{\phi_2}$ *intersects the diagonal* $T^*X \times \Delta_{T^*Y} \times T^*Z$ *transversally, where* Δ_{T^*Y} *is the diagonal of* $T^*Y \times T^*Y$.

(d) Given compact $K_X \subset\subset X$, $K_Z \subset\subset Z$, *the set of* y *for which* $a(x,y,\theta_1)b(y,z,\theta_2) \neq 0$ *for some* $(x,z,\theta_1,\theta_2) \in K_X \times K_Z \times \mathbb{R}^{N_1} \times \mathbb{R}^{N_2}$ *is compact.*

Then $A \circ B$ is a Fourier integral operator of order $m_1 + m_2$ (with nondegenerate phase), and its canonical relation is the composition of A and B's canonical relations.

Proof. The FIO composition theorem is a standard argument [1, 3]. We present part of it here that will be useful later, only sketching the remaining details. We refer the reader to the full proofs in the references.

First, it can be shown that conditions (a) and (d) imply the composition $A \circ B$ is defined as a linear operator $AB\colon C^\infty(Z) \to \mathcal{D}'(X)$. If K_A, K_B are the kernels, the idea of the proof is that the composition can be defined as an operator whose kernel is the restriction of $K_A \otimes K_B$ to the diagonal $T^*X \times \Delta_{T^*Y} \times T^*Z$, followed by pushforward by the projection onto $T^*X \times T^*Z$.

Now composing A and B, for $f \in C_0^\infty(Z)$,

$$ABf(x) = c \int\limits_{\mathbb{R}^{N_1}} \int\limits_{Y} \int\limits_{\mathbb{R}^{N_2}} \int\limits_{Z} e^{i[\phi_1(x,y,\theta_1)+\phi_2(y,z,\theta_2)]} a(x,y,\theta_1) b(y,z,\theta_2) f(z)\, dz\, d\theta_2\, dy\, d\theta_1. \quad (1.3.37)$$

(we let c be the factors of 2π, which are not essential for the argument.) By condition (c), the set of y for which $ab \neq 0$ for any θ_i and any $z \in \operatorname{supp} f$ is compact. Hence we can continuously expand the integral in y to all of $\mathbb{R}^n$ by setting it to zero for y outside of Y, then substitute $y = \mathfrak{y}/H$, where $H = |(\theta_1, \theta_2)|$. This makes $\mathfrak{y}$ a "frequency variable," so that the phase is homogeneous of degree 1 in $(\mathfrak{y}, \eta, \eta')$:

$$ABf(x) \;=\; c \int\limits_{\mathbb{R}^{N_1}} \int\limits_{\mathbb{R}^{n_Y}} \int\limits_{\mathbb{R}^{N_2}} \int\limits_{Z} e^{i[\phi_1(x,\mathfrak{y}/H,\theta_1)+\phi_2(\mathfrak{y}/H,z,\theta_2)]} a(x,y,\theta_1) b(y,z,\theta_2) f(z)\, dz\, d\theta_2\, dy\, d\theta_1. \quad (1.3.38)$$

Letting $\phi = \phi_1(x, \frac{\mathfrak{y}}{H}, \theta_1) + \phi_2(\frac{\mathfrak{y}}{H}, z, \theta_2)$, and setting $d_{\mathfrak{y},\eta,\eta'}\phi = 0$, we find the critical set for the new phase is

$$C_\phi = \{(x, \mathfrak{y}, z, \theta_1, \theta_2) \mid d_y\phi_1 = -d_y\phi_2,\ d_{\theta_i}\phi_i = 0\}. \quad (1.3.39)$$

As is, ab is not a symbol, since for instance $\partial_{\mathfrak{y}}^\alpha \partial_{\theta_1}^\beta a$ only decays as $O((1+|\theta_1|)^{m_1-|\beta|})$ and not as $O((1+|(\theta_1,\theta_2)|)^{m_1-|\beta|})$. Similarly, ϕ may not be smooth near $\theta_1 = 0$ and $\theta_2 = 0$. To handle both problems, we introduce a cutoff $\rho(\theta_1, \theta_2)$, homogeneous of order 0 in (θ_1, θ_2), such that $\rho = 0$ in a neighborhood of $\{\theta_1 = 0\} \cup \{\theta_2 = 0\}$. Having done that, we write AB

as a sum

$$ABf(z) = c\left[\int e^{i\phi} ab\rho \, dx \, d\theta_1 \, d\mathfrak{y} \, d\theta_2 + \int e^{i\phi} ab(1-\rho) \, dx \, d\theta_1 \, d\mathfrak{y} \, d\theta_2\right]. \tag{1.3.40}$$

We can check that $ab\rho$ does define a symbol. The phase ϕ is real, smooth on the support of $ab\rho$, and $d\phi \neq 0$ as a result of condition (b). Finally, condition (c) is required to ensure that the phase ϕ is nondegenerate. Hence, the first integral now defines an FIO.

Meanwhile, the kernel of the second integral can be shown to be C^∞, by a variation of the argument that an oscillatory integral is smooth if $d\phi \neq 0$ on the support of the amplitude (choosing a suitable first-order operator L such that $L\phi = 1$ and $L^k(ab)$ is $O((1+|(\eta_1, \eta_2)|)^{m_1+m_2-k})$). □

1.3.4 Fourier Integral Operators on Manifolds

So far we have defined and worked FIOs in local form, that is, as oscillatory integrals. It turns out, however, that is more natural to think of FIOs not as local integrals with phase and amplitude but as integral operators whose Schwartz kernels are distributions on a Lagrangian manifold $\Lambda' \subseteq T^*X$.

The following proposition reinforces this point:

Proposition 1.12. *Let $\phi_1(x, \theta_1)$, $\phi_2(x, \theta_2)$ be two nondegenerate phase functions parameterizing identical Lagrangians $\Lambda_{\phi_1} = \Lambda_{\phi_2} \subseteq T^*X$. If u is a (local) Fourier integral distribution associated with the phase ϕ_1 and an amplitude a_1, there exists an amplitude a_2 such that*

$$u(x) = \int e^{i\phi_2(x,\theta_2)} a_2(x, \theta_2) \, d\theta_2 \quad \mod C^\infty. \tag{1.3.41}$$

Proof Idea. [3] Let ρ be a bump function and consider the Fourier transform of the distributions $u = \rho \int e^{i\phi_1} a_1 \, d\theta_1$, $v = \rho \int e^{i\phi_2} a_2 \, d\theta_2$, and apply the method of stationary phase to write down an asymptotic development of $\widehat{\rho u}$ and $\widehat{\rho v}$. Equating the top terms yields the leading order term of a_2 and the lower order terms are found subsequently. □

As well, any Lagrangian submanifold of T^*X can be parameterized locally by a phase function:

Proposition 1.13. *Let $X \subseteq \mathbb{R}^n$, and $\Lambda \subseteq T^*X$. Then there exists a nondegenerate phase function $\phi(x, \theta)$ parameterizing Λ microlocally near any $\gamma \in T^*X$.*

Hence it makes sense to define a Fourier integral operator as associated with a Lagrangian rather than a certain phase, which allows us to define global Fourier integral operators as well as extending them to manifolds.

Definition 1.13. Let X be a manifold, and Λ a Lagrangian submanifold of T^*X. Then a Fourier integral operator A on X with canonical relation Λ' is a locally finite sum of local Fourier integral operators $A = \sum A_j$ whose Lagrangians Λ_j lie inside Λ.

BIBLIOGRAPHY

[1] J.J. Duistermaat, Fourier Integral Operators, Modern Birkhäuser Classics, doi:10.1007/978-0-8176-8108-1_2

[2] L. Hörmander, Analysis of Linear Partial Differential Operators, vols. III–IV. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 274–275. Springer-Verlag, Berlin, 1994.

[3] G. Uhlmann, Microlocal analysis class notes, 1996–1997.

[4] J.-F. Trèves, Introduction to Pseudodifferential and Fourier Integral Operators, University Series in Mathematics, Springer, 1980.

Chapter 2

SYNTHETIC APERTURE RADAR AND SINGULARITIES

2.1 Introduction

In synthetic aperture radar (SAR), a radar apparatus attached to an airplane or satellite travels along a given path, bouncing microwave radiation off the ground and measuring the echos. The goal is to form an image of ground reflectivity. A simplified mathematical model [26][1] for the measurements made in SAR, assuming an undirected beam and flat terrain, is the circular Radon transform

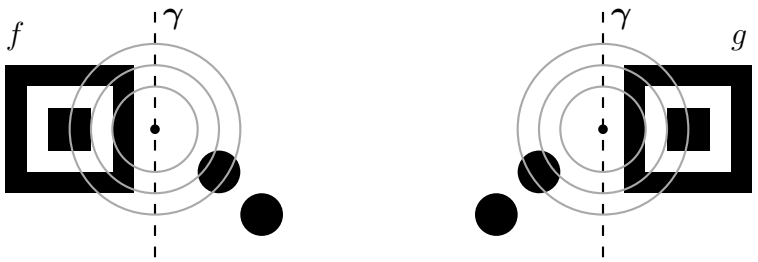
$$(R_\gamma f)(r,s) = \int_{\{|x-\gamma(s)|=r\}} f(x)\,\mathrm{d}\ell(x). \tag{2.1.1}$$

Here f is the ground reflectivity function (or distribution), $\gamma(s)$ is the flight path, a smooth curve in $\mathbb{R}^2$, and $x \in \mathbb{R}^2$; $d\ell$ is the Euclidean measure on the circle $\{|x-\gamma(s)| = r\}$. In other words, $R_\gamma f$ is the collection of line integrals of f over circles whose centers lie on γ. For a more sophisticated model, effects such as the antenna beam pattern or varying ground height can be included [12, 21, 23]. The inverse problem for SAR is recovering the image f from the data $R_\gamma f$, as much as is possible.

A simple example shows that this inverse problem cannot always be solved. Let γ be a straight line, and let f be any function; then let g be the reflection of f across this line. If

[1] See also [21], p.2

we look at the SAR data, then by symmetry $R_\gamma f$ and $R_\gamma g$ are identical:



So, if we are surveying the scene f and have the SAR data $R_\gamma f$, we will not be able to discern whether it came from f, or from g. In general, we *cannot* recover f from $R_\gamma f$. This leads to the question we want to address: for which γ can we recover f from $R_\gamma f$?

When f is a continuous, compactly-supported function, this problem was solved by Agranovsky and Quinto in [4]. They proved $R_\gamma f$ is injective if and only if γ is not contained in a straight line[2]. In other words, the inverse problem is only impossible for straight flight paths, at least when f is a smooth, compactly supported function. The injectivity of the circular Radon transform is a problem many authors have studied; a very small list of references is [2, 3, 5, 13, 14, 15, 18, 25]; [17] has many more. Beyond SAR, the transform has applications to photoacoustic and thermoacoustic tomography [3, 15], and other imaging problems ([18] has more examples).

Instead of recovering a continuous function f, we could drop the assumption that f is continuous and set ourselves the goal of recovering f's *singularities*. This is the problem we will consider. To be more precise, we let $\gamma(s)$ be a smooth curve, parameterized by arc length, and let f be a distribution (with $\gamma \cap \operatorname{singsupp} f = \emptyset$). We are now interested in the *microlocal injectivity question*: can we recover $\operatorname{WF}(f)$ from $\operatorname{WF}(R_\gamma f)$? This question was taken up by Stefanov and Uhlmann in [26]. Following that paper, we will use the tools of microlocal analysis, and view R_γ as a Fourier integral operator (FIO). Their paper is one of quite a few, including [1, 12, 21, 23, 22, 27], which have taken a microlocal approach to SAR.

[2]Their result was actually more general; in their paper they consider general circular transforms where the smooth curve γ is replaced by an arbitrary subset of $\mathbb{R}^2$.

The canonical transformation, χ, of R_γ is a 2–1 map, which reflects the left-right ambiguity we saw earlier. Given (x, ξ), suppose the line through x in the direction ξ intersects γ at $\gamma(s) = x + t\xi/|\xi|$. Then χ sends (x, ξ) to a singularity at $(r, s) = (|t|, s)$ with direction $(\rho, \sigma) = ((\operatorname{sgn} t)\,|\xi|, \dot{\gamma}(s) \cdot \xi)$. Now, χ also sends the *mirror image* of (x, ξ) with respect to the tangent line to γ at s to (r, s, ρ, σ) (figure 2.1).

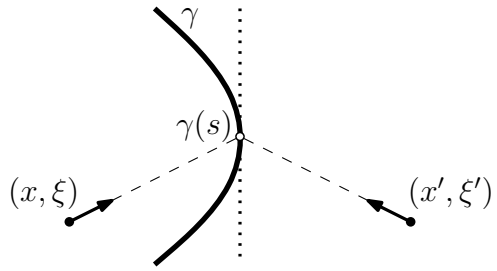


Figure 2.1: Mirror image singularities.

With this in mind, let us take a look at what it takes for microlocal injectivity to be false for γ, following [26]. Suppose we have an f that is singular, but $R_\gamma f$ is smooth. Choose any (x_0, ξ_0) in $\mathrm{WF}(f)$, and suppose the line through (x_0, ξ_0) crosses γ at $\gamma(s_1)$. Now χ maps (x_0, ξ_0) to some $\alpha \in T^*(\mathbb{R}_+ \times \gamma)$. But $R_\gamma f$ is smooth, and the only way $R_\gamma f$ can escape having a singularity at α is for this singularity to be cancelled by an opposite singularity at (x_0, ξ_0)'s mirror image. Call that mirror image (x_1, ξ_1). Now the line through (x_1, ξ_1) may intersect γ at another point, say $\gamma(s_2)$, in which case it must be cancelled at the mirror point (x_2, ξ_2), and so on (figure 2.2)[3].

In other words, the presence of a singularity at (x_0, ξ_0) will force mirror image singularities at other points. When γ is the boundary of a closed, strictly convex region, Stefanov and Uhlmann showed that the mirror images ((x_i, ξ_i), $i = 0, 1, 2, \ldots$ in our example) tend to infinity. Using this, they were able to prove that all the singularities in $\mathrm{WF}(f)$ one could

[3] Of course, the line through any of the (x_i, ξ_i) could intersect γ at more than two points. In general, if $(x, \xi) \in \mathrm{WF}(f)$, then at each point of intersection between γ and the line through (x, ξ), there must be an opposite singularity in f at the corresponding mirror image of (x, ξ).

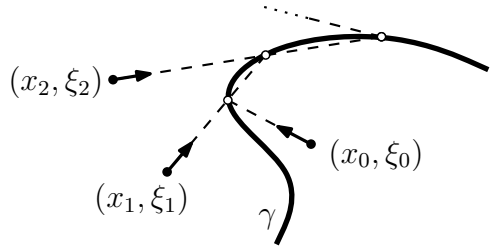


Figure 2.2: One singularity necessitates a sequence of cancelling singularities.

hope to recover can in fact be recovered, under the assumption that singsupp f is compact [26, thm. 2.3]. This compactness assumption they show is necessary as well as sufficient.

So we have a form of microlocal injectivity for closed γ bounding strictly convex regions. What happens for more general γ? The key feature of closed convex γ is that mirror images tend to infinity. If instead, the mirror images formed a loop, so at some point (x_n, ξ_n) returned to (x_0, ξ_0), we might be able to set up an f with compact singular support for which the singularities at $(x_0, \xi_0), \ldots, (x_{n-1}, \xi_{n-1})$ all cancel. Working locally, we replace γ by a sequence of curves $\gamma_1, \ldots, \gamma_n$, each representing γ near the points of intersection.

Then, we have two major questions we can ask about $\gamma_1, \ldots, \gamma_n$:

Discrete Question. *Is there an open set of covectors (x, ξ) that is preserved after reflecting across $\gamma_1, \ldots, \gamma_n$ (consecutively)?*

Continuous Question. *Is there a distribution f that is singular, but whose SAR data $R_{\gamma_1} f, \ldots, R_{\gamma_n} f$ is smooth?*

Let us say that the γ_j *allow discrete cancellation* if the answer to the first question is true, and that they *allow continuous cancellation* if the answer to the second question is true. (Here, "cancellation" is cancellation of singularities.) We investigate these two questions further in the following sections.

First, section 2.2 addresses the continuous question. By the argument with mirror points earlier, the γ_j must allow discrete cancellation if they allow continuous cancellation. The

main result is that any set of γ_j that allows discrete cancellation also allows continuous cancellation up to leading order, meaning $R_{\gamma_j} f$ can be one order smoother than it would be for arbitrary γ_j (corollary 2.3). In other words, if the γ_j allow discrete cancellation, there is the potential that objects can be hidden when the flight path traverses the γ_j.

Next, section 2.3 discusses the discrete question, giving the (limited) examples of γ_j we know of that allow discrete cancellation. In section 2.4, we take one of these example choices of γ_j, and develop a function f numerically that illustrates continuous cancellation.

2.2 Continuous Cancellation

2.2.1 Terminology

Let $\gamma_1, \dots, \gamma_n$ be smooth plane curves, parameterized with unit speed by $s_1, \dots, s_n$, with domains $I_j \subseteq \mathbb{R}$. Let $\operatorname{supp} f \subseteq \mathbb{R}^2 \setminus \gamma$. Define the SAR measurement operators

$$(R_{\gamma_j} f)(r, s_j) = \int\limits_{\{|x-\gamma_j(s_j)|=r\}} f(x)\, \mathrm{d}\ell(x) = \frac{1}{2\pi} \iint \mathrm{e}^{\mathrm{i}\alpha(|x-\gamma_j(s_j)|-r)} f(x)\, \mathrm{d}\alpha\, \mathrm{d}x. \tag{2.2.1}$$

The domain of $R_{\gamma_j} f$ is $(0, \infty) \times I_j$. The second form shows R_{γ_j} as an FIO; like the standard 2D Radon transform it has order $-\frac{1}{2}$. The condition $\operatorname{supp} f \subseteq \mathbb{R}^2 \setminus \gamma$ is necessary to ensure that the phase is smooth. R_{γ_j}'s canonical relation can be written: [26]

$$\Lambda' = \left\{ \left(r,\, s_j,\, x,\, \pm|\xi|\,,\, \xi \cdot \dot\gamma_j(s_j),\, \xi \right) : x = \gamma_j(s_j) \pm r \frac{\xi}{|\xi|},\, r > 0 \right\}. \tag{2.2.2}$$

For $(a, b) \in \mathring{T}^*\mathbb{R}^2 = T^*\mathbb{R}^2 \setminus \{0\}$, we will use "the line through (a, b)" to mean the line in $\mathbb{R}^2$ through a in direction b. Also, for a covector $\xi = (\xi^1, \xi^2)$, let $\xi^\perp = (-\xi^2, \xi^1)$ be its rotation counterclockwise by 90°. We say that a line through (x, ξ) intersects γ_j at $\gamma_j(s_j)$ *from the left* if $\dot\gamma_j(s_j)^\perp \cdot \xi < 0$, and *from the right* if $\dot\gamma_j(s_j)^\perp \cdot \xi > 0$. In both cases, (x, ξ) is said to be *visible* from γ_j at $\gamma_j(s_j)$. Conversely, if the line through (x, ξ) does not intersect γ_j, or intersects tangentially, (x, ξ) is not visible from γ_j.

Assume $W_0, \dots, W_n \subseteq \mathring{T}^*\mathbb{R}^2$ are conic open sets, such that for $j = 1, \dots, n$,

(a) For all $(x, \xi) \in W_{j-1}$, the line through (x, ξ) intersects γ_j exactly once from the left;

(b) For all $(x,\xi) \in W_j$, the line through (x,ξ) intersects γ_j exactly once from the right.

With these assumptions, we can define *discrete reflection maps* $\mathcal{C}_j : W_{j-1} \to \mathring{T}^*\mathbb{R}^2$, which send a covector to its mirror image. Given $(x,\xi) \in W_{j-1}$, let $\gamma_j(s_j)$ be the point of intersection of the line through (x,ξ) with γ_j, and let $r = (x - \gamma_j(s_j)) \cdot \xi$. The reflected ray will have direction $\xi' = \xi - 2(\xi \cdot \dot{\gamma}_j(s_j)^\perp)\dot{\gamma}_j(s_j)^\perp$. Then set $\mathcal{C}_j(x,\xi) = (\gamma_j(s_j) - r\xi'/|\xi'|,\, \xi')$.

Now (by shrinking the W_j if necessary) we can make one more assumption:

(c) $\mathcal{C}_j$ maps W_{j-1} bijectively onto W_j.

By [26], theorem 2.1, for each j there exists a unitary FIO U_j [4], whose canonical relation is the graph of $\mathcal{C}_j$, such that

$$R_{\gamma_j}(f_\mathrm{L} - f_\mathrm{R}) \in C^\infty \iff f_\mathrm{R} - U_j f_\mathrm{L} \in C^\infty \tag{2.2.3}$$

for any distributions $f_\mathrm{L} \in \mathcal{D}'(W_{j-1})$, $f_\mathrm{R} \in \mathcal{D}'(W)$. Here $\mathcal{D}'(W)$, for $W \subseteq T^*X$, represents $\{u \in \mathcal{D}'(X) \mid \mathrm{WF}(u) \subseteq W\}$. Equation 2.2.3 uniquely identifies U_j modulo smoothing operators; for, if $\tilde{U}_j$ also satisfies (2.2.3) then $(U_j - \tilde{U}_j)f_\mathrm{L} = (f_\mathrm{R} - \tilde{U}_j f_\mathrm{L}) - (f_\mathrm{R} - U_j f_\mathrm{L}) \in C^\infty$ for all f_L. Define U to be the composition $U_n \circ U_{n-1} \circ \cdots U_1$; the canonical relation of U is given by $\mathcal{C} = \mathcal{C}_n \circ \mathcal{C}_{n-1} \circ \cdots \circ \mathcal{C}_1$.

2.2.2 Main Results

Theorem 2.1. *If $\mathcal{C}$ is the identity, then U is a pseudodifferential operator with principal symbol equal to 1 on W_0.*

Corollary 2.2. *If $\mathcal{C} = \mathrm{Id}$, and $f \in H^s(\mathbb{R}^2)$, $\mathrm{WF}(f) \subseteq W_0$, then $Uf - f \in H^{s+1}(\mathbb{R}^2)$.*

Corollary 2.3. *Suppose $\mathcal{C} = \mathrm{Id}$ and n is even. Let $\gamma\colon I \to \mathbb{R}^2$ be a smooth curve containing $\gamma_1, \cdots, \gamma_n$ (meaning that I contains all the I_j and $\gamma_j = \gamma\big|_{I_j}$). Assume the line through each $(x,\xi) \in W_0 \cup \cdots \cup W_{n-1}$ intersects γ exactly twice. Then, if $g \in H^s(\mathbb{R}^2)$, $\mathrm{WF}(g) \subseteq W_0$, there*

[4] Note that this differs from the definition of U_j in [26] by a factor of -1.

exists $f \in H^s(\mathbb{R}^2)$ *extending* g *such that* $R_\gamma f \in H^{s+3/2}$ *and* $\mathrm{WF}(f-g) \cap (W_0 \setminus (W_1 \cup \cdots \cup W_{n-1})) = \emptyset$.

Theorem 2.1 and corollary 2.2 imply that, as long as the canonical relation $\mathcal{C}$ preserves an open conic set of singularities, reflecting across $\gamma_1, \ldots, \gamma_n$ preserves the magnitude of those singularities to leading order. In Section 2.3 we present several examples where $\mathcal{C}$ is the identity on an open set, so theorem 2.1 implies that for those examples the continuous reflection operator U is microlocally the identity, to leading order, as well. In fact, for all examples of $\mathcal{C} = \mathrm{Id}$ we know of, it is true that $U = \mathrm{Id}$ microlocally to all orders, which invites the question of whether this is true in general.

Corollary 2.3 says that we have microlocal noninjectivity (to leading order at least) for $R_\gamma f$ whenever $\mathcal{C} = \mathrm{Id}$. This is quite different from the result in [4], which implies R_γ is noninjective on $C_0^\infty(\mathbb{R}^2)$ only when γ is a straight line. Section 2.3 gives two examples of four-curve systems $\gamma_1, \ldots, \gamma_4$ for which $\mathcal{C} = \mathrm{Id}$ (and we have microlocal noninjectivity) for which R_γ is not injective on C_0^∞.

2.2.3 Supporting Results

For proving the main results, we will make use of the supporting results in this section.

To start, let us make some further definitions. Let $(x_0, \xi_0) \in W_0$, and for $j = 1, \ldots, n$, let $(x_j, \xi_j) = \mathcal{C}_j \circ \mathcal{C}_{j-1} \cdots \circ \mathcal{C}_1(x_0, \xi_0)$. Let $z_j = \gamma_j(s_j)$ be the point where the line through (x_{j-1}, ξ_{j-1}) intersects γ_j; let c_j be the curvature of γ_j at that z_j, and θ_j the angle from $\dot\gamma_j(s_j)$ clockwise to ξ_{j-1}. Let $q_j = -2c_j \csc\theta_j$. Also, define $z_0 = x_0$, $z_{n+1} = x_n$, and let $w_j = (z_{j+1} - z_j) \cdot \xi_j / |\xi_j|$ $(0 \le j \le n)$. The $z_j, q_j, c_j, s_j, w_j, \theta_j$ are functions of (x_0, ξ_0) but we will usually write e.g. z_j instead of $z_j(x_0, \xi_0)$.

Now, we will write down a formula for the Jacobian $D\mathcal{C}(x_0, \xi_0)$, but first we need to settle coordinate systems. Let $\xi = (R\cos\alpha, R\sin\alpha)$. Define coordinates $(\delta u, \delta v, \delta R, \delta\alpha)$ on $T_{(x,\xi)}\mathring{T}^*\mathbb{R}^2$ as follows: for $(\delta x, \delta R, \delta\alpha) \in T_{(x,\xi)}\mathring{T}^*\mathbb{R}^2$, let

$$\delta x = \delta u \frac{\xi}{|\xi|} + \delta v \frac{\xi^\perp}{|\xi^\perp|}, \tag{2.2.4}$$

To be more down-to-earth, all we are doing here is rotating the x coordinate system, replacing it by (u, v) coordinates, with the u coordinate parallel to ξ and the v coordinate perpendicular to ξ.

Proposition 2.4 (Jacobian of $\mathcal{C}$)**.** *In $(\delta u, \delta v, \delta\alpha, \delta R)$ coordinates, the Jacobian of $\mathcal{C}$ is*

$$D\mathcal{C} = \begin{bmatrix} 1 & & & \\ & (-1)^n & & \\ & & (-1)^n & \\ & & & 1 \end{bmatrix} + (-1)^n \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & D & E & 0 \\ 0 & C & \tilde{D} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \tag{2.2.5}$$

$D\mathcal{C}$ has determinant 1. Here $E = E_{0,n}$, $D = D_{0,n}$, $\tilde{D} = \tilde{D}_{0,n}$, $C = C_{0,n}$, where

$$\begin{aligned} E_{a,b} &= \textstyle\sum_{j_0,\dots,j_{2l}} \quad w_{j_0} q_{j_1} w_{j_2} q_{j_3} \cdots q_{j_{2l-1}} w_{j_{2l}} \\ D_{a,b} &= \textstyle\sum_{j_1,\dots,j_{2l}} \quad q_{j_1} w_{j_2} q_{j_3} \cdots q_{j_{2l-1}} w_{j_{2l}} \\ \tilde{D}_{a,b} &= \textstyle\sum_{j_0,\dots,j_{2l-1}} \quad w_{j_0} q_{j_1} w_{j_2} \cdots w_{j_{2l-2}} q_{j_{2l-1}} \\ C_{a,b} &= \textstyle\sum_{j_1,\dots,j_{2l-1}} \quad q_{j_1} w_{j_2} q_{j_3} \cdots w_{j_{2l-2}} q_{j_{2l-1}} \\ &\qquad (a \le j_0 < j_1 \le j_2 < j_3 \le \cdots < j_{2l-1} \le j_{2l} \le b,\ l \ge 1) \end{aligned} \tag{2.2.6}$$

Proposition 2.4 can be proved by directly computing the Jacobians for each reflection, in $(u_j, v_j, \alpha_j, R_j)$ coordinates, and multiplying them together. The condition $\det D\mathcal{C} = 1$ follows from $\mathcal{C}$ being a canonical transformation.

The symmetry in (2.2.5) and (2.2.6) is interesting in itself, but we can also connect $D = D_{0,n}$ with the continuous reflection operator U, using the next proposition.

First, we define the following *no caustics* assumption. Let π_x, π_ξ be the projections of $\mathring{T}^*\mathbb{R}^2$ onto its first and second factors, respectively.

Assumption (No caustics)**.** For all $x_j \in \pi_x(W_j)$ and $\xi_0 \in \pi_\xi(W_0)$, there is only one x_0 such that $\mathcal{C}(x_0, \xi_0)$ lies over x_j $(1 \le j \le n)$.

This assumption rules out cases such as the parabola with x_1 at the focus and ξ_0 perpendicular to the directrix, where the mirror image of (x, ξ_0) is over x_1 for all x on the directrix.

Proposition 2.5.

(a) U is an FIO of order 0 with canonical relation $\mathcal{C}$ and principal symbol 1, with respect to trivialization of the half-density and Maslov bundles described in [19, after prop. 25.3.3].

(b) Furthermore, under the no-caustics assumption, U can be represented as

$$Uf(x_n) = \frac{1}{(2\pi)^2} \int \mathrm{e}^{\mathrm{i}\phi_n(x_n,\xi_0)} a(x_n,\xi_0) \hat{f}(\xi_0)\, \mathrm{d}\xi_0, \tag{2.2.7}$$

with generating function

$$\phi_n(x_n,\xi_0) = x_0 \cdot \xi_0 + (w_0 + \cdots + w_n)\,|\xi_0| \tag{2.2.8}$$

and $a(x_n,\xi_0) \sim \sum_{j=0}^{\infty} a_{-j}$ is an amplitude whose terms satisfy the transport equations

$$\frac{\partial a_{-j}}{\partial w_n} = -\frac{\Delta_{x_n}\phi_n}{2|\xi_0|} a_{-j} + \frac{\mathrm{i}}{2|\xi_0|} \Delta_{x_n} a_{-j+1} \qquad (j > 0). \tag{2.2.9}$$

(Δ_{x_n} is the Laplacian with respect to the x_n variables.) The leading term a_0 also satisfies the transport equation if we define $a_1 = 0$, and has the formula

$$a_0(x_n,\xi_0) = (1+D)^{-1/2} = \left((-1)^n \frac{\partial v_n}{\partial v_0}\right)^{-1/2} \tag{2.2.10}$$

D is defined in (2.2.6).

To construct the FIO representation of U, we will use a slightly different definition for the reflection operator. First, we define the wave solution operator Λ_γ. Given $g \in \mathcal{E}'(\mathbb{R}^2)$, let u solve the constant-speed wave equation

$$\begin{cases} (\partial_t^2 - \Delta_x) u(t,x) = 0, & (t,x) \in \mathbb{R}_+ \times \mathbb{R}^2, \\ u(0,x) = 0, & x \in \mathbb{R}^2, \\ \partial_t u(0,x) = g, & x \in \mathbb{R}^2. \end{cases} \tag{2.2.11}$$

and define $\Lambda_\gamma g \in \mathcal{D}'((0,\infty)\times\gamma)$ by the restriction $(\Lambda_\gamma g)(t,s) = u(t,\gamma(s))$. In other words, u is the solution to the wave equation with initial conditions $(0,g)$, and $\Lambda_\gamma g$ is its values on γ for all time.

By [26, thm. 2.1], we can write

$$U_j = \Lambda_{j,\mathrm{R}}^{-1}\Lambda_{j,\mathrm{L}}, \tag{2.2.12}$$

where $\Lambda_{j,\mathrm{L}}$, $\Lambda_{j,\mathrm{R}}$ are the restrictions of Λ_{γ_j} to distributions whose singularities only intersect γ_j from the left or right, respectively. The codomain of $\Lambda_{j,\mathrm{L}}$ and $\Lambda_{j,\mathrm{R}}$ is chosen so that the wavefront relations of these operators are bijections; see [26] for the details.

We can equivalently characterize U_j as follows: for any $f_\mathrm{L} \in \mathcal{D}'(W_{j-1})$, $f_\mathrm{R} \in \mathcal{D}'(W_j)$,

$$\Lambda_{\gamma_j}(f_\mathrm{L} - f_\mathrm{R}) \in C^\infty \iff f_\mathrm{R} - U_j f_\mathrm{L} \in C^\infty. \tag{2.2.13}$$

Sketch of Proof for Proposition 2.5.

Part (a). The U_j are FIOs of order 0 whose canonical relations are the graphs of $\mathcal{C}_j$, so their composition $U = U_n \circ \cdots \circ U_1$ is a well-defined FIO of order 0 associated with the canonical transformation $\mathcal{C}_n \circ \cdots \circ \mathcal{C}_1$.

To find the principal symbol, we first consider a single reflection U_j. Let $E\colon \mathcal{D}'(\mathbb{R}^2) \to \mathcal{D}'(\mathbb{R}\times\mathbb{R}^2)$ be the solution operator to the wave equation (2.2.11); i.e., Eg is the solution to the unit-speed wave equation with Cauchy data $(0,g)$. E can be written as a sum of FIOs E_+, E_-:

$$E = E_+ + E_- = \int \mathrm{e}^{\mathrm{i}(x\cdot\xi + t|\xi|)}\frac{\hat f(\xi)}{2i|\xi|}\,\mathrm{d}\xi + \int \mathrm{e}^{\mathrm{i}(x\cdot\xi - t|\xi|)}\frac{-\hat f(\xi)}{2i|\xi|}\,\mathrm{d}\xi. \tag{2.2.14}$$

Write R for the FIO restricting $\mathbb{R}\times\mathbb{R}^2$ to $(0,\infty)\times\gamma$. Then $\Lambda = \Lambda_+ + \Lambda_-$, where $\Lambda_\pm = R\circ E_\pm$. Now $\sigma_0(E)$ is by symmetry independent of x and $\xi/|\xi|$, so it is a function of $|\xi|$, t, τ; hence $\sigma_0(\Lambda_\pm)$ is a function of t, s, τ, σ. Let $\Lambda_{\mathrm{L},\pm}$, $\Lambda_{\mathrm{R},\pm}$ be the restrictions of $\Lambda_\pm$ to $\mathcal{D}'(W_0)$ and $\mathcal{D}'(W_1)$, respectively. Then $\Lambda_\mathrm{L} = \Lambda_{\mathrm{L},+} + \Lambda_{\mathrm{L},-}$, $\Lambda_\mathrm{R} = \Lambda_{\mathrm{R},+} + \Lambda_{\mathrm{R},-}$. Define sets Σ_+, Σ_- by

$$\Sigma_\pm = \{(x,\xi) \mid x \pm a\xi \in \gamma \text{ for some } a > 0\}. \tag{2.2.15}$$

By (2.2.2), the images of the canonical relations of Λ_+, Λ_- are disjoint ($\rho < 0$ in the first; $\rho > 0$ in the second). Hence $U_j = \Lambda_{\mathrm{R},+}^{-1}\Lambda_{\mathrm{L},+}$ on Σ_+ and $U_j = \Lambda_{\mathrm{R},-}^{-1}\Lambda_{\mathrm{L},-}$ on Σ_- modulo smoothing operators. As $\Lambda_{\mathrm{R},+}$ is elliptic, $\sigma_0(\Lambda_{\mathrm{R},+}^{-1}) = \sigma_0(\Lambda_{\mathrm{R},+})^{-1}$. Both $\Lambda_{\mathrm{L},+}$ and $\Lambda_{\mathrm{R},+}$ are restrictions of Λ_+ so on Σ_+

$$\begin{aligned}\sigma_0(U_j) = \sigma_0(\Lambda_{\mathrm{R},+}^{-1})\sigma_0(\Lambda_{\mathrm{L},+}) \quad &= \tfrac{\sigma_0(\Lambda_{\mathrm{L},+})(t,s,\tau,\sigma)}{\sigma_0(\Lambda_{\mathrm{R},+})(t,s,\tau,\sigma)} \\ &= \tfrac{\sigma_0(\Lambda_+)(t,s,\tau,\sigma)}{\sigma_0(\Lambda_+)(t,s,\tau,\sigma)} = 1. \end{aligned} \tag{2.2.16}$$

For this step, it was important for $\sigma_0(\Lambda_\pm)$ to be independent of x. The same holds on Σ_-, and hence each U_j has principal symbol 1. This implies U has principal symbol 1 too.

Part (b). First, assume $n = 1$, and let $\Lambda = \Lambda_{\gamma_1}$, $U = U_1$, etc. The idea is to write U in terms of the sum of two FIOs, then solve for the amplitudes and phases and verify this results in the correct operator. To start, suppose (formally) that $f = F_\xi(x) = \mathrm{e}^{\mathrm{i}x\cdot\xi}$ is a plane wave, where ξ is fixed. We want to find Uf such that $\Lambda(Uf) = \Lambda f$; by (2.2.13), this characterizes U.

The solution to the wave equation with Cauchy data $(0, F_\xi)$ is

$$u_0(t,x) = EF_\xi(t,x) = \frac{1}{2\mathrm{i}|\xi|}\mathrm{e}^{\mathrm{i}(x\cdot\xi+t|\xi|)} - \frac{1}{2\mathrm{i}|\xi|}\mathrm{e}^{\mathrm{i}(x\cdot\xi-t|\xi|)}. \tag{2.2.17}$$

We then try to find a new solution of the wave equation with the form

$$u_1(t,x) = \mathrm{e}^{\mathrm{i}(\phi_+(x,\xi)+t|\xi|)}b_+(t,x,\xi) + \mathrm{e}^{\mathrm{i}(\phi_-(x,\xi)-t|\xi|)}b_-(t,x,\xi), \tag{2.2.18}$$

such that $u_1\big|_{t=0} = 0$ and $u_0 = u_1$ for $x \in \gamma_1$. Here ϕ_+, ϕ_- are phase functions, which we can choose to be identical, and b_+, b_- are amplitudes.

Write out the amplitudes b_+, b_- in asymptotic expansions: $b_\pm \sim \sum_{j=0}^\infty b_{\pm,-j}$, where $b_{\pm,-j} = O(|\xi|^{-j})$. Requiring (2.2.18) to be a solution to the wave equation leads to an eikonal equation for $\phi_+ = \phi_-$ and transport equations for the amplitude terms similar to (2.2.9). The eikonal equation has two solutions. One is the original, $x\cdot\xi$, which would lead to $U = \mathrm{Id}$; this does satisfy $U\Lambda f = \Lambda f$, but is not the operator we want. Hence, the other solution,

which turns out to be the phase function claimed in (2.2.8), must be correct. It also turns out that $b_- = -b_+$, so we drop the signs and write $b = b_+$.

To get our proposed UF_ξ we take a time derivative of (2.2.18) and set $t = 0$, yielding $UF_\xi = \mathrm{e}^{\mathrm{i}\phi(x,\xi)} 2\mathrm{i}|\xi| b(0, x, \xi)$. With this done, we can use linearity to write down Uf formally as an FIO, as in (2.2.7):

$$Uf = \frac{1}{(2\pi)^2} \int \hat{f}(\xi) UF_\xi \, \mathrm{d}\xi = \frac{1}{(2\pi)^2} \int \mathrm{e}^{\mathrm{i}\phi(x,\xi)} 2\mathrm{i}|\xi| \, b(0, x, \xi) \hat{f}(\xi), d\xi. \tag{2.2.19}$$

Then we set $a = 2\mathrm{i}|\xi|b$. By construction, $\Lambda(Uf) = \Lambda f$, so to verify the validity of our formula, all that is left to do is to check that $\mathrm{WF}(U)$ maps W_0 to W_1. We can do this by computing the canonical relation of U and checking that it is equal to $\mathcal{C}$.

For $n > 1$ we repeat this process with $\gamma_1, \ldots, \gamma_n$ successively. Starting with $f = F_\xi$ and u_0 as before, we look for a solution u_1 of the form (2.2.18) matching u_0 on γ_1, then look for u_2 of the same form (2.2.18) matching u_1 on γ_2, and so on to u_n. At each step, we have to solve an eikonal equation for the phase ϕ_j and we obtain transport equations for the amplitudes, which lead to (2.2.9).

To obtain the amplitude formula (2.2.10) we start by deriving

$$\Delta_{x_n} \phi_n = |\xi| \cdot \frac{C_n}{C_n w_n - \sin\theta_n}. \tag{2.2.20}$$

where the *twist* C_n is defined by $C_n = (\partial\alpha_n/\partial v_0)/(\partial s_n/\partial v_0) = C/(\partial s_n/\partial v_0)$. We solve the transport equation for a_0 exactly, then use a recursive formula for C_n and some algebra to achieve (2.2.10). Note that we can compute that the principal symbol is 1 using this formula for a_0 and the principal symbol formulas in [19, after prop. 25.3.3]. □

2.2.4 Proofs of Main Results

Proof of theorem 2.1. This is a direct consequence of Proposition 2.5. As $\mathcal{C}$ is the identity map, U is a ΨDO, so according to Proposition 2.5 it has principal symbol 1 with respect to the trivialization of the half-density and Maslov bundles used for pseudodifferential operators.

□

Proof of corollary 2.2. Let $\chi(x,\xi)$ be a C^∞ bump function equal to 1 on $\mathrm{WF}(f)$ and 0 on $\mathring{T}^*\mathbb{R}^2\backslash W_0$, and let $P=\chi(x,D)$ be the pseudodifferential operator with χ as its symbol. Then $(U-\mathrm{Id})f=(\mathrm{Id}-P)(U-\mathrm{Id})f+P(U-\mathrm{Id})f$. The first term is smooth, since $\mathrm{WF}(Uf-f)\subseteq \mathrm{WF}(f)$ and $\mathrm{Char}(\mathrm{Id}-P)\cap\mathrm{WF}(f)=\emptyset$.

In the second term, the principal symbol of $P(U-\mathrm{Id})$ is the product of χ and the principal symbol a_0-1 of $U-\mathrm{Id}$. Since at each $(x,\xi)\in\mathring{T}^*\mathbb{R}^2$ either χ or a_0-1 is zero, $P(U-\mathrm{Id})$ has zero principal symbol, so it is a pseudodifferential operator of order -1. Hence $P(U-\mathrm{Id})f\in H^{s+1}$, which makes $(U-\mathrm{Id})f\in H^{s+1}$. □

Proof of corollary 2.3. Define $g_0=g$, $g_j=U_jU_{j-1}\cdots U_1 g$ for $1\le j<n$, then set $f=g_0-g_1+g_2-g_3+\cdots-g_{n-1}$. Our goal is to show that $R_\gamma f\in H^{s+3/2}$, which we will do microlocally.

Let $\tilde{W}_0,\ldots,\tilde{W}_{n-1}$ be the images of $W_0,\ldots,W_{n-1}$ under the canonical relation Λ_{R_γ} of R_γ. By assumption, the line through any (x,ξ) in any W_j intersects γ once from the left and once from the right. Hence we can decompose each $\tilde{W}_j$ into a disjoint union $\tilde{W}_{j,L}\cup\tilde{W}_{j,R}$, where $\tilde{W}_{j,L}$, $\tilde{W}_{j,R}$ are the images due to intersections from the left and right, respectively. Assumption (c) on the W_j (page 30) states $\mathcal{C}_j$ maps W_{j-1} bijectively onto W_j, and then assumptions (a–b) imply that $\tilde{W}_{j,L}=\tilde{W}_{j+1,R}$, $j=0,\ldots,n-1$. Also, since $\mathcal{C}=\mathcal{C}_n\circ\cdots\circ\mathcal{C}_1=\mathrm{Id}$, assumption (c) also implies $W_n=W_0$, so $W_{n-1,L}=W_{0,R}$.

Let $Y_j=\tilde{W}_{j,R}$, and let us agree to read all indices modulo n. Because $\tilde{W}_{j,L},\tilde{W}_{j,R}$ are disjoint,

$$Y_j\cap Y_{j+1}=\emptyset,\qquad j=1,\ldots,n. \tag{2.2.21}$$

Note that by assumption (b), $Y_j\subseteq T^*(\mathbb{R}_+\times I_j)$, so $R_\gamma f=R_{\gamma_j}f$ on Y_j.

Now, suppose $\alpha\in T^*(\mathbb{R}_+\times I)$, and let $J\subseteq\{1,\ldots,n\}$ be the set of indices j such that $\alpha\in Y_j$. By (2.2.21), $j+1$, $j-1$ are not in J if $j\in J$. We can write $R_\gamma f$ in the form

$$R_\gamma f=\sum_{j\in J}R_\gamma\left((-1)^j\left(g_j-g_{j-1}\right)\right)+\sum_{\{j:\,j,\,j+1\notin J\}}(-1)^jR_\gamma g_j \tag{2.2.22}$$

Near α, the second sum of (2.2.22) is smooth, since $\mathrm{WF}(R_\gamma g_j)$ is a closed subset of the open set $\tilde{W}_j=Y_j\cup Y_{j+1}$, which does not contain α. We now turn to the first sum of (2.2.22). If

$j < n$, then

$$R_\gamma (g_j - g_{j-1}) = R_\gamma (U_j g_{j-1} - g_{j-1}) . \tag{2.2.23}$$

By definition of U_j (2.2.3), the right hand side is smooth. On the other hand, if $j = n$, then

$$\begin{aligned} R_\gamma (g_j - g_{j-1}) &= R_\gamma (g_0 - g_{n-1}) \\ &= R_\gamma (g_0 - U_n g_{n-1}) + R_\gamma (U_n g_{n-1} - g_{n-1}) \\ &= R_\gamma (g_0 - U g_0) + R_{\gamma_n} (U_n g_{n-1} - g_{n-1}) . \end{aligned} \tag{2.2.24}$$

The second term on the right-hand side is again smooth by (2.2.3). As for the first term, corollary 2.2 implies that $g_0 - U g_0 \in H^{s+1}$.

Now, R_γ is an FIO of order $-\frac{1}{2}$ and its canonical relation is locally a canonical graph, so R_γ maps $H^{s+1}(\mathbb{R}^2)$ to $H^{s+3/2}(\mathbb{R}_+ \times I)$ [19, thm. 25.3.1]. So we have shown that $R_\gamma f \in H^{s+3/2}$ in a neighborhood of α. Hence, any $\alpha \in T^*(\mathbb{R}_+ \times I)$ has a neighborhood on which $R_\gamma f$ is in $H^{s+3/2}$, which implies $R_\gamma f \in H^{s+3/2}(\mathbb{R}_+ \times I)$.

Finally, we observe that since $\mathrm{WF}(g_i) \subseteq W_i$, we know $f - g = g_1 + \cdots + g_{n-1}$ is smooth outside of $W_1 \cup \cdots \cup W_{n-1}$, and in particular, $\mathrm{WF}(f - g) \cap (W_0 \setminus (W_1 \cup \cdots \cup W_{n-1})) = \emptyset$. We also claimed $f \in H^s$: since the U_i are FIOs associated with canonical graphs, they map H^s to H^s, and because $g_0 \in H^s$, it follows all the g_i are in H^s too, and therefore f is also. □

2.2.5 Algebraic Cancellation Conditions

Before considering the discrete cancellation question, we have one more theorem, based on the description of the Jacobian in Proposition 2.4. Recall from (2.2.6) that E is a polynomial in the w_j and q_j, and recall q_j depends on the curvature of γ_j and the angle of intersection with γ_j, and that w_j is the distance from one curve intersection point to the next, or to the beginning and ending points.

Theorem 2.6. *If $E \equiv 0$ on W_0, then U is the pullback by some rigid motion $x \mapsto Mx + d$, modulo an FIO of order -1.*

Theorem 2.6 says that the pointwise algebraic condition $E = 0$ is enough to show that Uf is to leading order the composition of a rotation and translation of f. This is an avenue for potential work on the cancellation problem in the future..

Proof. First, we show that $E \equiv 0$ implies $\mathcal{C}$ is a rigid motion, a map of the form $(x, \xi) \mapsto (Mx + d, M\xi)$, where $M \in \mathbf{O}(2)$, $d \in \mathbb{R}^2$. This will imply $Uf(Mx + d) = f(x) + Jf(x)$, where J is some ΨDO of order -1.

$\mathcal{C}$ is a rigid motion. Let $(x_0, \xi_0) \in W_0$. Suppose we perturb x_0 by $t\xi_0$, where t is small; this does not affect the intersection points $z_1, \dots, z_n$, nor $w_1, \dots, w_{n-1}$. On the other hand, $w_0(x_0 + t\xi_0, \xi_0) = w_0(x_0, \xi_0) - t$, while $w_n(x_0 + t\xi_0, \xi_0) = w_n(x_0, \xi_0) + t$. Hence $w_0 + w_n$ is equal to a constant, say K.

Recall the definition of $E_{a,b}$ from (2.2.6): $E_{a,b} = \sum w_{j_0} q_{j_1} w_{j_2} \cdots w_{j_{2l-2}} q_{j_{2l-1}} w_{j_{2l}}$, where the sum is taken over $a \le j_0 < j_1 \le j_2 < \cdots \le j_{2l-2} < j_{2l-1} \le j_{2l} \le b$. We can separate out the sum $E = E_{0,n}$ into four parts, depending on whether $j_0 = 0$ and whether $j_{2l} = n$, and similarly for D and $\tilde{D}$, to rewrite them as follows:

$$\begin{aligned} E &= w_0 w_n C + w_0 D_{0,n-1} + w_n \tilde{D}_{1,n} + E_{1,n-1} \\ D &= w_n C + D_{0,n-1} \\ \tilde{D} &= w_0 C + \tilde{D}_{1,n} \end{aligned} \tag{2.2.25}$$

Substituting $w_0 = K - w_n$ and rearranging gives

$$E = -w_n^2 C + w_n \Big[KC - D_{0,n-1} + \tilde{D}_{1,n} \Big] + \Big[K D_{0,n-1} + E_{1,n-1} \Big]. \tag{2.2.26}$$

By perturbing x_0 in the direction ξ_0 (while staying in W_0), we vary w_n while preserving the other coefficients in E. Since $E \equiv 0$, the coefficients of w_n^2 and w_n in E must be identically zero as well. This gives $C \equiv 0$, and $D_{0,n-1} \equiv \tilde{D}_{1,n}$. By (2.2.25), $D = \tilde{D}$.

Proposition 2.4 gives $1 = \det D\mathcal{C} = (1+D)(1+\tilde{D}) - CE$, implying $(1+D)^2 = 1$. Therefore $D\mathcal{C}$ is diagonal with entries $(1, \pm 1, \pm 1, 1)$; this holds for all $(x_0, \xi_0) \in W_0$. Recalling the $(\delta u, \delta v, \delta\alpha, \delta R)$ coordinates on $T\mathring{T}^*\mathbb{R}^2$ (see (2.2.4)), this implies $\mathcal{C}(x, \xi) = (Mx + d, M\xi)$, for some $M \in \mathbf{O}(2)$ and $d \in \mathbb{R}^2$.

U is pullback by a rigid motion. We can write U in the standard form

$$Uf(x_n) = \frac{1}{(2\pi)^2} \int \mathrm{e}^{\mathrm{i}\phi_n(x_n,\xi_0)} a(x_n,\xi_0)\hat{f}(\xi_0)\,\mathrm{d}\xi_0, \tag{2.2.27}$$

which has canonical relation

$$\left\{ (x_n,\xi_n;x_0,\xi_0) \;\middle|\; x_0 = \frac{\partial\phi_n}{\partial\xi_0},\ \xi_n = \frac{\partial\phi_n}{\partial x_n} \right\}. \tag{2.2.28}$$

This is the graph of $\mathcal{C}$. If $S(x_n,\xi_0) = M^{-1}(x_n - d)\cdot\xi_0$, then $\partial S/\partial\xi_0 = \partial\phi_n/\partial\xi_0$, $\partial S/\partial x_n = \partial\phi_n/\partial x_n$. This implies $\phi_n - S$ is a constant, which must be zero since ϕ_n, S are both homogeneous in ξ_0. So $S = \phi_n$ and we have

$$\begin{aligned} Uf(Mx_0+d) &= \tfrac{1}{(2\pi)^2}\textstyle\int \mathrm{e}^{\mathrm{i}S(Mx_0+d,\xi_0)} a(Mx_0+d,\xi_0)\hat{f}(\xi_0)\,\mathrm{d}\xi_0, \\ &= \tfrac{1}{(2\pi)^2}\textstyle\int \mathrm{e}^{\mathrm{i}x_0\cdot\xi_0} a(Mx_0+d,\xi_0)\hat{f}(\xi_0)\,\mathrm{d}\xi_0. \end{aligned} \tag{2.2.29}$$

By proposition 2.5, U has principal symbol 1, which implies $a = 1$ modulo symbols of order -1. Hence $Uf(Mx_0+d) = \mathrm{Id}$ modulo a ΨDO of order -1. □

2.3 Discrete Cancellation

So far, we know of only two (easy) examples where discrete cancellation happens, that is, where $\mathcal{C} = \mathrm{Id}$. In both cases, the reflection operator U is not only microlocally identity plus a lower order error term, as theorem 2.1 implies, but in fact $U = \mathrm{Id}$ in a conic open set, which invites the question of whether this is true in general.

$n = 2$: With two curves γ_1, γ_2, it is not hard to show that the only setup where singularities cancel is the trivial one with $\gamma_1 = \gamma_2$. This is a result of a uniqueness lemma (Lemma 2.10).

$n = 4$: With four curves, we know of two families of non-trivial examples.

In the first family of examples (a,a'), γ_2 and γ_4 are segments lying along a common line, and γ_1 and γ_3 are any curves that are mirror images with respect to this line. (Later, in figure 2.7, we will see an example of a singularities cancelling for a set-up of this type.)

In the second family $(\mathrm{b}, \mathrm{b}')$, $\gamma_1, \ldots, \gamma_4$ are segments of lines that intersect in a common point, and the angle from γ_1 to γ_2 must be identical to the angle from γ_4 to γ_3. The intersection point may be infinite (γ_j parallel), in which angles are replaced by distances between the γ_j.

The next proposition says that if one of the γ_j is a line segment, these configurations are the only ones where $\mathcal{C} = \mathrm{Id}$. Recall that z_j is the point where the line through (x_{j-1}, ξ_{j-1}) intersects γ_j.

Proposition 2.7. *Assume $\gamma_1, \ldots, \gamma_4$ do not intersect. If $\mathcal{C} = \mathrm{Id}$ in a neighborhood W_0 of (x_0, ξ_0), and γ_2 is a line segment, then γ_4 is a line segment in a neighborhood of z_4, and locally near z_1, z_3 either*

(a) γ_2 and γ_4 are collinear and γ_1, γ_3 are mirror image curves (as in figure 2.3a), or

(b) γ_1, γ_3 are also line segments (as in figure 2.3b), with the angle from γ_1 to γ_2 equal to that from γ_4 to γ_3.

(c) γ_1, γ_3 are also line segments, the γ_j are parallel, and the distance from γ_1 to γ_2 is equal to that from γ_4 to γ_3 (as in figure 2.3b, but with parallel lines).

The assumption that $\gamma_1, \ldots, \gamma_4$ do not intersect rules out the trivial case where $\gamma_1 = \gamma_4$ and $\gamma_2 = \gamma_3$. Note that there may be other arrangements of four curves where singularities cancel if none of the curves are straight, but so far we have not found any.

To prove Proposition 2.7, we will use the following lemmas:

Lemma 2.8. *Let $\ell_1, \ldots, \ell_4$ be four lines, and let $\mathcal{C}_{\ell_j} \colon \mathring{T}^*\mathbb{R}^2 \to \mathring{T}^*\mathbb{R}^2$ be the associated reflection maps. Then $\mathcal{C}_\ell = \mathcal{C}_{\ell_4} \circ \cdots \circ \mathcal{C}_{\ell_1}$ has a fixed point if and only if it is identity, and if and only if $\ell_1, \ldots, \ell_4$ have a common point of intersection p in $\mathbb{RP}^2$ (including $\mathbb{R}^2$ in $\mathbb{RP}^2$ in the standard way, $(x, y) \mapsto (x : y : 1)$), and*

(a) if p is finite ($p \in \mathbb{R}^2$), the angle from ℓ_1 to ℓ_2 equals the angle from ℓ_4 to ℓ_3;

(b) if p is infinite ($p \in \mathbb{RP}^2 \setminus \mathbb{R}^2$), the signed distance from ℓ_1 to ℓ_2 equals that from ℓ_4 to ℓ_3.

Proof. Identify $\mathbb{R}^2$ as $\mathbb{C}$ and describe each line ℓ_i as a (non-unique) pair of complex numbers (a_i, v_i), where $\ell_i = a_i + \mathbb{R}v_i$ and $|v_i| = 1$. A quick calculation shows

$$\mathcal{C}_{\ell_i} = T^{-1} \circ R \circ T = (v_i^2\overline{(z - a_i)} + a_i, v_i^2\overline{\zeta}). \tag{2.3.1}$$

Composing the $\mathcal{C}_{\ell_i}$ and simplifying gives

$$\begin{aligned}\mathcal{C}_\ell \quad = \Big(&\tfrac{v_4^2v_2^2}{v_3^2v_1^2}(z - a_1) + \tfrac{v_4^2v_2^2}{v_3^2}\overline{(a_1 - a_2)} + \tfrac{v_4^2}{v_3^2}(a_2 - a_3) + v_4^2\overline{(a_3 - a_4)} + a_4, \\ &\tfrac{v_4^2v_2^2}{v_3^2v_1^2}\zeta\Big).\end{aligned} \tag{2.3.2}$$

If $\mathcal{C}_\ell$ fixes one (z, ζ), then (2.3.2) implies that $v_4^2v_2^2/v_3^2v_1^2 = 1$, and $\mathcal{C}_\ell$ is actually the identity. This proves the lemma's first "if and only if," so we proceed to the second part.

For convenience, call $v_4^2v_2^2/v_3^2v_1^2 = 1$ the *angle condition*; it states that the angle between ℓ_1 and ℓ_2 equals that between ℓ_4 and ℓ_3, modulo π. Now, we address the second "only if" and suppose $\mathcal{C}_\ell = \mathrm{Id}$; there are two cases.

ℓ_1, ℓ_2 not parallel. If ℓ_1, ℓ_2 are not parallel, then by the angle condition neither are ℓ_3 and ℓ_4. By changing coordinates, translate the intersection point of ℓ_1 and ℓ_2 to the origin, and rotate so that $v_3 = 1$. We can choose $a_1 = a_2 = 0$, and choose $a_3 = a_4$ to be the intersection point of ℓ_3 and ℓ_4. Then (2.3.2) implies $-v_4^2a_3 + a_3 = 0$, and therefore $a_3 = 0$, since $v_4 \neq \pm v_3 = \pm 1$. That says all ℓ_i pass through the origin, and we are in case (a).

ℓ_1, ℓ_2 parallel. On the other hand, if ℓ_1, ℓ_2 are parallel, (2.3.1) implies the composition $\mathcal{C}_{\ell_2} \circ \mathcal{C}_{\ell_1}$ is a translation:

$$\mathcal{C}_{\ell_2} \circ \mathcal{C}_{\ell_1}(z, \zeta) \quad = \Big(z + 2[iv_2 \operatorname{Re}(iv_2\overline{(a_2 - a_1)})],\ \zeta\Big) \tag{2.3.3}$$

The term $d(\ell_1, \ell_2) = \operatorname{Re}(iv_2\overline{(a_2 - a_1)})$ is the signed distance from a_1 to a_2. The angle condition implies ℓ_3 and ℓ_4 are likewise parallel, so we get an analogous formula to (2.3.3) for ℓ_3,

ℓ_4. Combining all the reflections,

$$\mathcal{C}_\ell(z,\zeta) = \Big(z + 2iv_2 d(\ell_1,\ell_2) + 2iv_4 d(\ell_3,\ell_4),\ \zeta\Big). \tag{2.3.4}$$

If $d(\ell_1,\ell_2) \neq 0$, we must have $v_2 = \pm v_4$, so the ℓ_i are all parallel, and $v_2 d(\ell_1,\ell_2) + v_4 d(\ell_3,\ell_4) = 0$. Hence, we are in case (b). On the other hand, if $d(\ell_1,\ell_2) = 0$ then $d(\ell_3,\ell_4) = 0$ also, and $\ell_1 = \ell_2$, $\ell_3 = \ell_4$; the ℓ_i have a common point of intersection and case (a) holds. This completes the "only if" direction.

"If" direction. For the "if" direction, when case (a) holds we can choose $a_1 = a_2 = a_3 = a_4$ and (2.3.2) together with the angle condition implies that $\mathcal{C}_\ell = \mathrm{Id}$. If case (b) holds, (2.3.4) implies $\mathcal{C}_\ell = \mathrm{Id}$. □

For the next lemma, recall that W_0 is the domain of $\mathcal{C}$, $z_i = \gamma_i(s_i)$ is the point where γ_i and the line through (x_{i-1},ξ_{i-1}) intersect, θ_i is the angle of intersection, c_i is the curvature of γ_i there, $q_i = -2c_i \csc\theta_i$, and finally $w_0 = |z_1 - x_0|$, $w_1 = |z_2 - z_1|$ (see figure 2.4).

Lemma 2.9. *Let $\gamma_1,\dots,\gamma_4$ be smooth curves, and assume $\mathcal{C}_{\gamma_4} \circ \dots \circ \mathcal{C}_{\gamma_1} = \mathrm{Id}$ on W_0. For $(x_0,\xi_0) \in W_0$, let $\ell_i = \ell_i(x_0,\xi_0)$ be the tangent line to γ_i at z_i. If the common point of intersection of $\ell_1,\dots,\ell_4$ is $p \in \mathbb{R}^2$ (the ℓ_i do not intersect at infinity), and $u_i = |z_i - p|$,*

$$q_1 u_1^2 = q_3 u_3^2, \qquad q_2 u_2^2 = q_4 u_4^2. \tag{2.3.5}$$

Here W_0, z_i, and q_i are defined as before.

Proof. We first focus on γ_1, γ_2. Consider perturbing (x_0,ξ_0) in order to vary θ_1 while preserving z_1. To be precise, define $(x(\alpha),\xi(\alpha))$ to be the point such that $x(\alpha) + w_1(x_0,\xi_0)\xi(\alpha) = \gamma_1(s_1(x_0,\xi_0))$ and $\theta_1(x(\alpha),\xi(\alpha)) = \alpha$. For α sufficiently close to $\theta_1(x_0,\xi_0)$, the $\ell_i(x(\alpha),\xi(\alpha))$ will not be parallel and therefore have a common intersection point.

We will arrive at (2.3.5) by computing $du_1/d\alpha$ in two different ways. By the law of sines, $u_1/\sin\alpha = w_1/\sin(\pi - \alpha - \theta_2)$. Using implicit differentiation, we compute

$$\frac{dw_1}{d\alpha} \qquad = w_1 \cot\theta_2$$

$$\frac{d\theta_2}{d\alpha} = c_2 w_1 \csc\theta_2 - 1 \tag{2.3.6}$$

Differentiating u_1, plugging in the derivatives (2.3.6) and simplifying gives

$$\frac{du_1}{d\alpha} = c_2 w_1^2 \csc^2(\theta_2 + \alpha) \csc\theta_2 \sin\alpha. \tag{2.3.7}$$

Using the law of sines again, we have

$$\frac{du_1}{d\alpha} = c_2 u_2^2 \csc\alpha \csc\theta_2. \tag{2.3.8}$$

Now, we can do exactly the same with γ_4 in place of γ_2, so

$$c_4 u_4^2 \csc\alpha \csc\theta_4 = \frac{du_1}{d\alpha} = c_2 u_2^2 \csc\alpha \csc\theta_2. \tag{2.3.9}$$

Multiplying both sides by $-2\sin\alpha$ and recalling $q_i = -2c_i \csc\theta_i$, we have

$$q_4 u_4^2 = q_2 u_2^2. \tag{2.3.10}$$

By symmetry, $q_1 u_1^2 = q_3 u_3^2$. □

Finally, we have a small uniqueness lemma:

Lemma 2.10. *Let $\mathcal{B} : W \to \mathring{T}^*\mathbb{R}^2$ be a continuous map, $W \subset \mathring{T}^*\mathbb{R}^2$. Suppose there are two smooth curves γ and η such that $\mathcal{B} = \mathcal{C}_\gamma\big|_W = \mathcal{C}_\eta\big|_W$ (in particular $\mathcal{C}_\gamma$, $\mathcal{C}_\eta$ must be defined on W). Suppose $\gamma\big|_J$ is a segment from γ such that each point of $\gamma\big|_J$ is on at least one line through a covector $(x,\xi) \in W$. Then η also contains $\gamma\big|_J$.*

In other words, given $\mathcal{B} : W \to \mathring{T}^*\mathbb{R}^2$, there is essentially at most one curve γ for which $\mathcal{B} = \mathcal{C}_\gamma$. Another curve η may differ from γ only where lines through covectors in W do not intersect it. We will use this in the proof of Proposition 2.7 to show that if three of the γ_i are known, the last one is determined as well.

Proof of Lemma 2.10. Let J be as in the lemma, and $p \in \gamma\big|_J$. Choose $(x,\xi) \in W$ such that p is on the line m_1 through (x,ξ), and let $(x',\xi') = \mathcal{B}(x,\xi)$. Let m_2 be the line equidistant

from x and x', or if $x = x'$, the line through them x bisecting the angle from ξ to ξ'. Since (x', ξ') is the reflection of (x, ξ), p must be on m_2. By assumption (a) in the definition of $\mathcal{C}$ (see page 30), m_1 intersects m_2 transversally, and p is their intersection point.

Now, since $\mathcal{C}_\eta$ is defined at (x, ξ) there is a unique $p' \in \eta$ on the line through (x, ξ), and $p' \in m_1 \cap m_2 = \{p\}$. Hence $p = p'$, finishing the lemma. □

Proof of Proposition 2.7. Define $\ell_i = \ell_i(x_0, \xi_0)$ to be the tangent line to γ_i at z_i, as in lemma 10. For each $(x_0, \xi_0) \in W_0$, since the reflection $\mathcal{C}_{\ell_4(x_0,\xi_0)} \circ \cdots \circ \mathcal{C}_{\ell_1(x_0,\xi_0)}$ fixes (x_0, ξ_0), the ℓ_i have necessarily a common point of intersection $p = p(x_0, \xi_0) \in \mathbb{RP}^2$.

γ_4 is locally a line segment. First, assume p is finite ($p \in \mathbb{R}^2$). Since γ_2 is a line segment, we have $q_2 = 0$. By lemma 2.9, this implies that for any $(x_0', \xi_0') \in W_0$, either $q_4 = 0$ or $u_4 = 0$. By continuity, there is a neighborhood $V \subseteq W_0$ of (x_0, ξ_0) on which either $q_4 = 0$ or $u_4 = 0$ identically.

We can think of s_4 as a map $s_4 \colon W_0 \to I_4$, which sends a singularity to the s_4 value at the intersection point z_4. It can be checked that $\mathrm{d}s_4$ has full rank, so s_4 is an open map; in particular, $S_4 = s_4(V)$ is open.

If $q_4 = 0$ on V, then the curvature of $\gamma_4\big|_{S_4}$ is identically zero, so $\gamma_4\big|_{S_4}$ is a line segment. Conversely, if $u_4 = 0$ on V, then $p(x_0', \xi_0')$ lies on $\gamma_4\big|_{S_4}$ for all $(x_0', \xi_0') \in V$. However, $p(x_0', \xi_0')$ also lies on the tangent line $\ell_2(x_0', \xi_0')$ to γ_2, and $\ell_2(x_0', \xi_0')$ does not depend on (x_0', ξ_0'), since γ_2 is a line segment. Therefore $\gamma_4\big|_{S_4}$ is contained in the line ℓ_2.

Pinning down γ_1, γ_3. We know γ_2 and γ_4 are locally line segments near z_2 and z_4. Since we are working locally, we can restrict γ_4 to a neighborhood of z_4 where it is a line segment. There are two options now.

First, suppose γ_2 and γ_4 lie on a common line ℓ. Let $\tilde{\gamma}_3$ be the reflection of γ_3 across ℓ. Then $\mathcal{C}_{\tilde{\gamma}_3} = \mathcal{C}_\ell \circ \mathcal{C}_{\gamma_3} \circ \mathcal{C}_\ell$. But, since $\mathcal{C}_{\gamma_4} \circ \cdots \circ \mathcal{C}_{\gamma_1}$ is locally the identity, it is also true that $\mathcal{C}_{\gamma_1} = \mathcal{C}_\ell \circ \mathcal{C}_{\gamma_3} \circ \mathcal{C}_\ell$ (since $\mathcal{C}_{\gamma_2} = \mathcal{C}_{\gamma_4} = \mathcal{C}_\ell$). By the uniqueness lemma, γ_1 coincides locally with the reflection $\tilde{\gamma}_3$ of γ_3, and this is case (a).

Now suppose γ_2 and γ_4 are not on a common line. Since p is finite, γ_2 and γ_4 cannot be parallel, and ℓ_2, ℓ_4 have a single point of intersection p. Hence, near z_1 every tangent line to γ_1 passes through the fixed point p, showing that γ_1 is locally a straight line. Similarly for γ_3; so we are in case (b). The angle condition in Lemma 2.8 implies that the γ_i satisfy the angle condition for case (b).

p infinite. Now, suppose p is infinite ($p \in \mathbb{RP}^2 \setminus \mathbb{R}^2$). Let $t_4 = (x_0 - z_4) \cdot \xi_0/|\xi_0|$. (Note $|t_4| = |x_0 - z_4|$ and the sign of t_4 is positive iff (x_0, ξ_0) points away from γ_4.) Then we can perturb (x_0, ξ_0) slightly; let $\xi(t)$ be the rotation of ξ_0 counterclockwise by t radians, and let $x(t) = z_4 + t_4\xi(t)/|\xi_0|$. Note that $z_4 = z_4(x(t), \xi(t))$ is unaffected by this perturbation. For sufficiently small t, $(x(t), \xi(t))$ lies within W_0. If p is finite for any such t, we can apply the argument for finite p to show that γ_4 is locally a line segment at z_4. Otherwise, we have an open set of t values for which $p(x(t), \xi(t))$ is infinite, so that the tangent lines $\ell_i(x(t), \xi(t))$ are all parallel to $\ell_4(x(t), \xi(t)) = \ell_4(x_0, \xi_0)$. Next, we can check the location of these tangent lines. A calculation shows that ds_1/dt and ds_3/dt are given by:

$$\frac{ds_1}{dt} = -\csc\theta_1\,(z_4 - z_1) \cdot \frac{\xi(t)}{|\xi_0|}, \qquad \frac{ds_3}{dt} = -\csc\theta_3\,(z_4 - z_3) \cdot \frac{\xi_r(t)}{|\xi_0|}, \tag{2.3.11}$$

where $\xi_r(t)$ is the reflection of $\xi(t)$ across ℓ_4. Since the γ_j do not intersect, these derivatives are nonzero, and s_1, s_3 change with t. Letting t vary, this shows the tangent lines to γ_1 and γ_3 in some neighborhoods of s_1 and s_3 are all parallel to the fixed line $\ell_4(x_0, \xi_0)$. Therefore γ_1 and γ_3 are locally line segments. Now, let $\tilde{\gamma}_4$ be the line ℓ_4, parameterized as a smooth curve. The composition $\tilde{\mathcal{C}} = \mathcal{C}_{\tilde{\gamma}_4} \circ \mathcal{C}_{\gamma_3} \circ \mathcal{C}_{\gamma_2} \circ \mathcal{C}_{\gamma_1}$ preserves (x_0, ξ_0). So, by Lemma 2.8 (used locally), $\tilde{\mathcal{C}}$ preserves a neighborhood of (x_0, ξ_0). Then the uniqueness Lemma 2.10 implies that γ_4 coincides with the line $\tilde{\gamma}_4$ near z_4, and we are in case (c). The distance condition in Lemma 2.8 implies that the γ_i satisfy the distance condition of case (c). □

2.4 Numerical Experiments

Our goal in this section is to produce a numerical example of singularities cancelling, following the results in section 2.2 and using one of the 4-curve setups in section 2.3. Specifically, we want a singular image f such that $R_{\gamma_1}f, \ldots, R_{\gamma_4}f$ are smooth. Throughout this section, "smooth" and "singular" are necessarily subjective terms.

The key algorithmic step is to compute the reflection operator U for a given curve γ numerically. The basis for the approach described here is (2.2.12), which factors U as $\Lambda_{\mathrm{R}}^{-1}\Lambda_{\mathrm{L}}$. Intuitively, this translates as solving the wave equation forward, restricting to γ, and then running the process in reverse.

As a completely different approach to computing U, we could use a generic FIO-computing algorithm such as those described in [8, 9, 10], applied to the representation of U in Proposition 2.5. An upcoming paper will illustrate using the algorithm in [10].

2.4.1 Algorithm description

To implement (2.2.12), we need methods for computing Λ_{L} and $\Lambda_{\mathrm{R}}^{-1}$.

Computing Λ_{L}***.*** Given input g, we solve the wave equation with initial conditions $(0, g)$ using a finite difference method, and restrict to a discretized version of γ to find $\Lambda_{\mathrm{L}}g$. A spectral method with a sufficiently large spatial grid is another possibility for this step. The only significant issue is the need to truncate the domain $\mathbb{R}_+ \times \mathbb{R}^2$ in space and time. To avoid waves bouncing from the edges of the spatial grid, as would occur with Dirichlet or Neumann boundary conditions, we use *perfectly matched layers* (PMLs), introduced by Bérenger [6]. We implement Grote and Sim's algorithm [16] for the wave equation with PMLs. In the time domain, we cut off the calculations at a finite time T, choosing T somewhat arbitrarily, but large enough so that all the visible singularities of g have reached γ by time T.

Computing $\Lambda_{\mathrm{R}}^{-1}$***.*** This is the trickier part. To compute $\Lambda_{\mathrm{R}}^{-1}$, we refer to Proposition 3.3 in [26], which says that if γ is a smooth curve and the boundary of an open set $\Omega \subseteq \mathbb{R}^2$, the

singular support of g is in a compact subset $K \subset \Omega$, and T is large enough, then we can form a left inverse for Λ by backprojecting the boundary data. Namely, let u be a solution to the wave equation

$$\begin{cases} (\partial_t^2 - \Delta_x)u = 0, & (t,x) \in [0,T] \times \Omega, \\ u(T,x) = 0, & x \in \Omega, \\ \partial_t u(T,x) = 0, & x \in \Omega, \\ u(t,\gamma(s)) = \chi(t) \cdot (\Lambda g)(t,s), & t \in [0,T],\ \forall s. \end{cases} \tag{2.4.1}$$

Here $\chi(t)$ is a smooth cutoff function, equal to 0 for $t > T$ and 1 for t less than some T_0 satisfying $T > T_0 > \max\{|x-y| : x \in \partial\Omega,\ y \in K\}$. As $(\Lambda g)(T,s)$ is nonzero in general, the cutoff χ enforces compatibility between the boundary conditions and the final conditions.

With u defined this way, $g - \partial_t u\big|_{t=0}$ is smooth; so, the map taking Λg to $\partial_t u\big|_{t=0}$ is a microlocal left inverse for Λ. Note that the visible singularities of g are all to one side of γ.

This method is only valid when γ is a *closed* curve but for arbitrary curves we have Proposition 3.2 in [26], which says that for a distribution $h(t,s)$ (with some restrictions on $\mathrm{WF}(h)$),

$$\Lambda_{\mathrm{R}}^{-1} h = 2\,\partial_t u_{\mathrm{R}}\big|_{t=0}, \tag{2.4.2}$$

where u_{R} is the "incoming" solution to the wave equation on $\mathbb{R}^2$ satisfying $u_{\mathrm{R}}(t,\gamma(s)) = h(t,s)$. Here, an *incoming* solution is one whose singularities lying on $\mathbb{R}_+ \times \gamma$ propagate to the left of γ as t increases.

We can simulate finding $\Lambda_{\mathrm{R}}^{-1} h$ numerically with a process similar in spirit to (2.4.1) and (2.4.2) that gives reasonable results. First, we replace Ω in (2.4.1) by $\mathbb{R}^2$ and use a finite difference method again to solve the wave equation, this time running time in reverse. Then, we recover the reflection by taking a time derivative at $t = 0$, yielding the *backprojection operator* $M\colon h \mapsto 2\,\partial_t u(0,x)$. $M\Lambda_{\mathrm{L}} g$ includes singularities on both sides of γ, so we finish by multiplying by a spatial cutoff $\psi(x)$ to restrict to the right side of γ.

Note that if γ were closed, our new algorithm reduces to the reconstruction in (2.4.1), except for the factor of 2. The reason for the factor of 2 for non-closed curves is that singu-

Reflection algorithm 1 Approximate $Ug = \Lambda_R^{-1}\Lambda_L g$

- Solve the wave equation (2.2.11) using a finite difference method and PMLs, yielding $h = \Lambda_L g$.
- Multiply $h(t,s)$ by a smooth time cutoff $\chi(t)$.
- Solve the wave equation (2.4.1), with $\Omega = \mathbb{R}^2$ and time running in reverse, again using finite differences and PMLs, yielding $u(t,x)$.
- Compute reconstructed image $(Mh)(x) = 2\,\partial_t u(0,x)$.
- Multiply Mh by smooth spatial cutoff $\psi(x)$ to restrict it to the right of γ.

larities from g propagate along rays in opposite directions and if $\mathrm{WF}(g) \subset W_0$ only one of these rays encounters γ. Loosely then, half the magnitude of the singularity is reconstructed and we compensate through multiplying by 2.

2.4.2 Example of reflection algorithm

Figure 2.5(a–e) illustrates the steps of the reflection algorithm, applied to an image g with several circles, and the parabolic curve γ_1. The final time is $T = 2$.

Let us denote by $\bar{g} = \psi M \Lambda g$ the approximate reflection calculated by the algorithm. If $\bar{g}$ were the *exact* reflection of g across γ_1 (that is, $\bar{g} = U_1 g$), then $R_{\gamma_1}(g - \bar{g})$ should be smooth. This is tested in figures 2.5g and 2.5h, which show the SAR measured data from γ_1 for both $g - \bar{g}$ and the original image g. $R_{\gamma_1}(g - \bar{g})$ is relatively smooth; the singularities in the SAR data for g (2.5g) are largely canceled after subtracting the approximate reflection $\bar{g}$ (2.5h). The final image (2.5i) shows the results of backprojecting the boundary data of $g - \bar{g}$. The singularities that were present in the original g are now mostly cancelled in the boundary data, and the backprojection $M\Lambda(g - \bar{g})$ is much smoother than $M\Lambda g$.

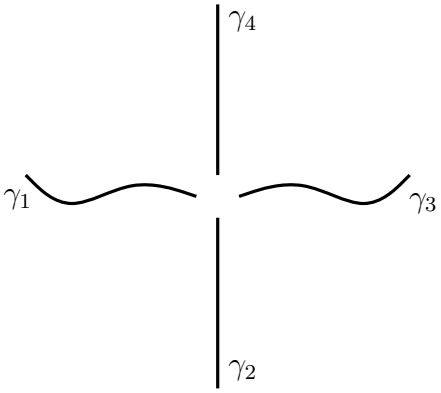
2.4.3 Hiding singularities

Finally, we consider the four curve setup $\gamma_1, \ldots, \gamma_4$ in figure 2.6 and attempt to construct a singular f whose singularities are "hidden" from the γ_j, meaning all the $R_{\gamma_j} f$ are smooth.

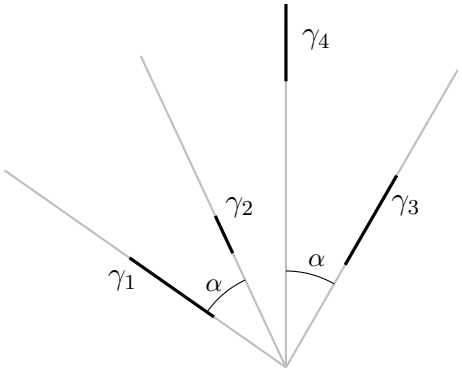
The four curves roughly divide the plane into four quadrants: I–IV. Suppose the γ_j are each parameterized so that γ_j' points inward, toward the origin, and that γ_2, γ_4 lie on the y-axis. Let U_j be the reflection operator for γ_j, $j = 1, \ldots, 4$.

Now, let $f \in \mathcal{D}'(\mathbb{R}^2)$, and let $f_1, \ldots, f_4$ be the restrictions of this unknown f to quadrants I–IV. Given f_1, equation (2.2.3) implies that $f_4 = -U_4 f_1$, $f_3 = -U_3 f_4$, and $f_2 = -U_2 f_3$, in order for all the SAR data $R_{\gamma_j} f$ to be smooth. The reflection operator for the straight line segments γ_2, γ_4 is just reflection about the y-axis, so an alternative method to generate f is to let $f_4 = -U_1 f_1$ (performing a simple horizontal reflection), then reflect $f_1 + f_4$ across $\gamma_1 \cup \gamma_3$ using our reflection algorithm. This approach has the advantage of allowing f_1 to have singularities that intersect γ_3 as well.

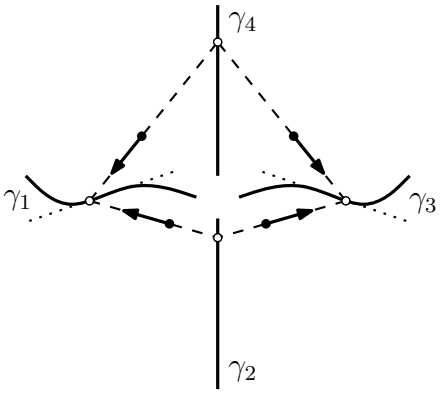
Figure 2.7a shows an f generated this way where f_1 is the image with circles used earlier. The simulated SAR measurements for f along two of the curves, γ_1 and γ_2, are shown in figures 2.7b and 2.7c. $R_{\gamma_1} f$ is fairly smooth except near the ends of the curve (top and bottom of figure 2.7b), and $R_{\gamma_2} f$ is essentially zero, unsurprisingly (2.7c). As another test, we apply M (with the union of the γ_j) to the SAR data Λf in figure 2.7d. The magnitude of the singularities are greatly reduced since the boundary data Λf is nearly smooth. Figure 2.7d is an indication of the image that we would be able to reconstruct from the measured SAR data if the ground reflectivity function were f. As promised, the circle pattern in quadrant I is mostly hidden.



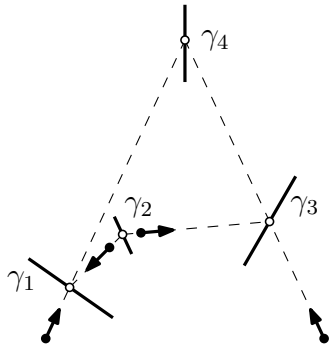
(a) Two line segments, two mirror curves



(b) Four line segments; the γ_j are the dark segments within the gray lines.



(a′) Example set of cancelling singularities



(b′) Example set of cancelling singularities

Figure 2.3: Four-curve set-ups where singularities can cancel.

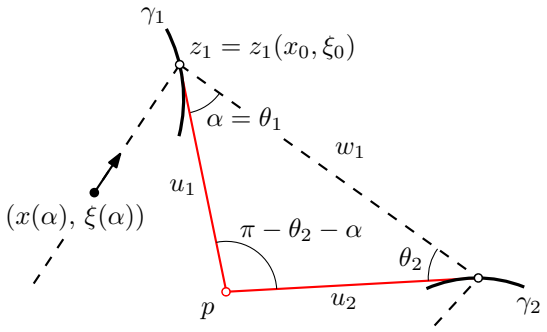
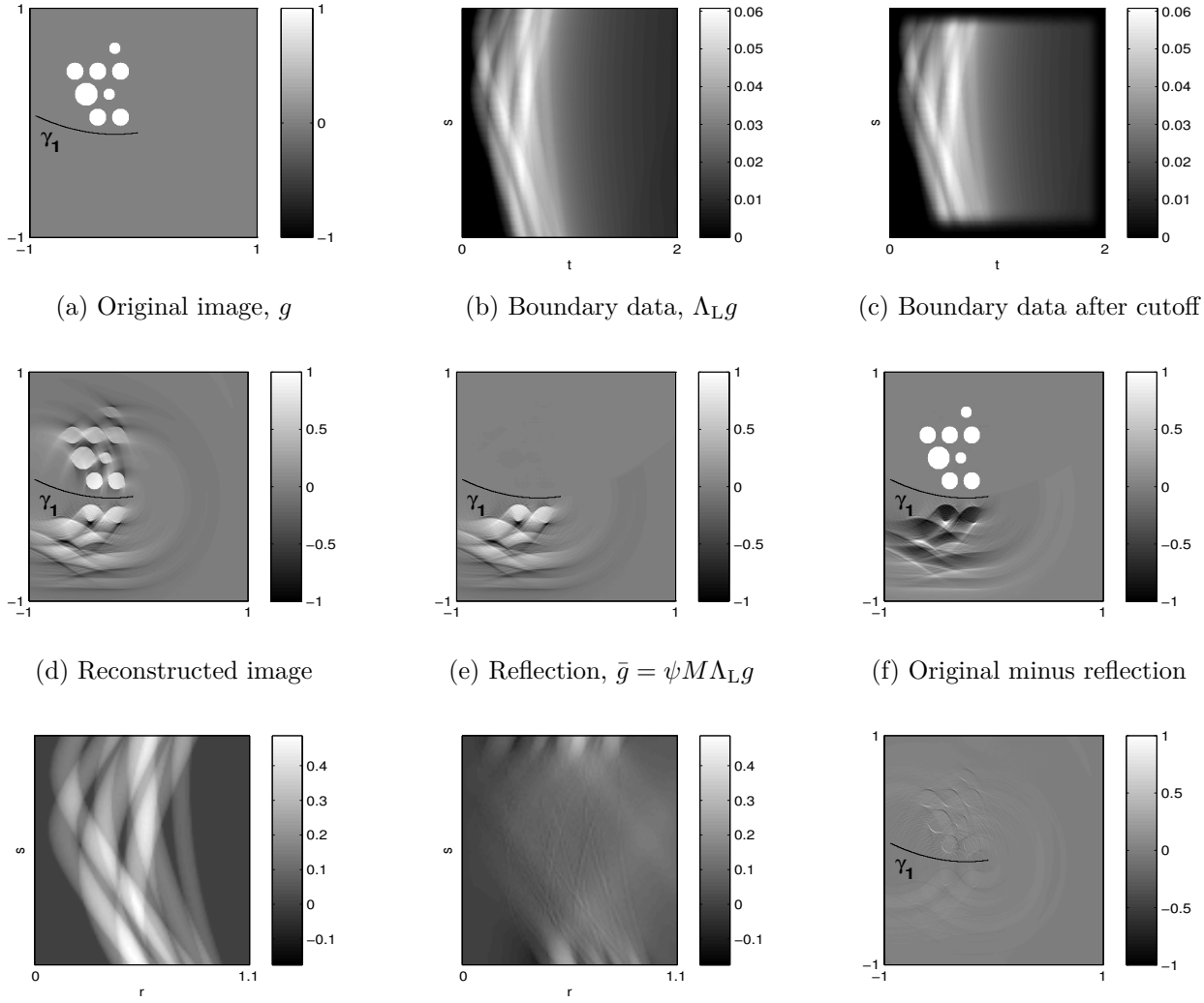


Figure 2.4: Setup for Lemma 2.9.



(a) Original image, g

(b) Boundary data, $\Lambda_{\mathrm{L}} g$

(c) Boundary data after cutoff

(d) Reconstructed image

(e) Reflection, $\bar{g} = \psi M \Lambda_{\mathrm{L}} g$

(f) Original minus reflection

(g) SAR data for original, $R_{\gamma_1} g$

(h) SAR data for original minus reflection

(i) Backprojected image $M\Lambda G$

Figure 2.5: Reflection algorithm in progress.

Original image (a). The wave equation is solved for $t \in [0, 2]$ with Cauchy data $(0, g)$; boundary data is measured (b) and smoothly cut off (c). The wave equation is solved with time running in reverse from $t = 2$ to $t = 0$ with the cut-off boundary data, generating (d). Multiplying by a cutoff gives an approximate reflection (e). Subtracting the reflection from original (f); the singularities of the original and reflection should cancel. Compared to SAR data for original (g), the SAR data for new image is close to zero (h), showing the singularities in f nearly cancel. A second backprojection step recovers little of the image (i).

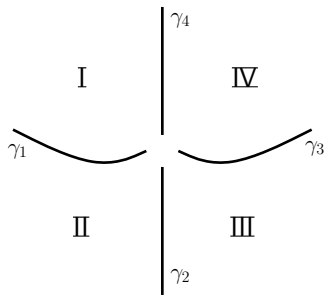
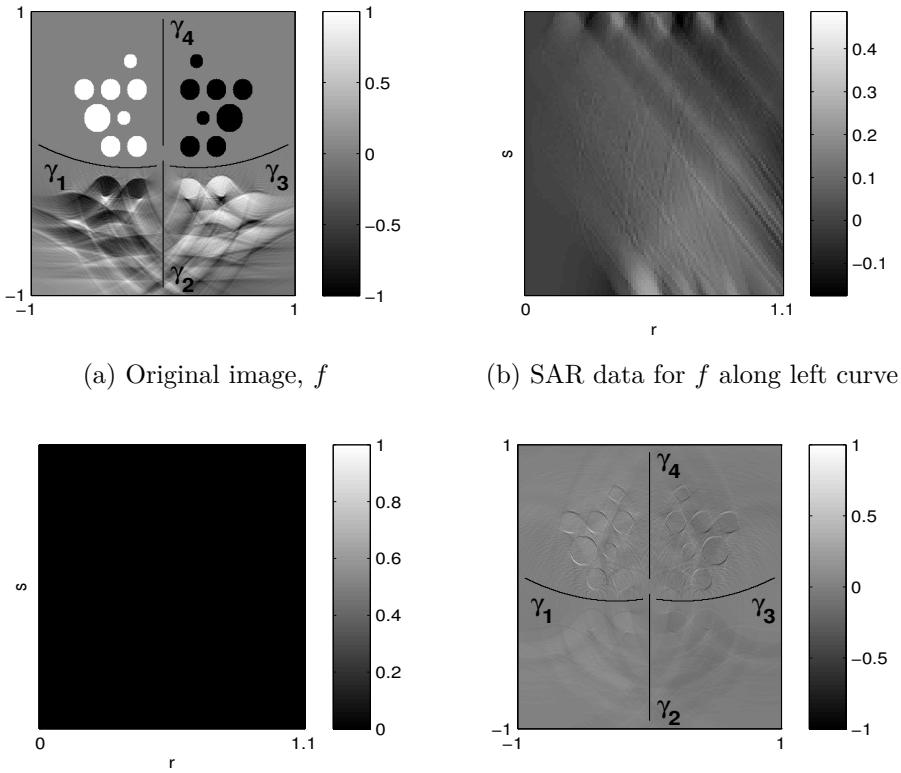


Figure 2.6: Four-curve setup.



(a) Original image, f

(b) SAR data for f along left curve

(c) SAR data for f along bottom curve

(d) Backprojected image, $M\Lambda f$

Figure 2.7: Approximate cancellation of singularities with four-curve setup.

The function f in (a) was created by the backprojection algorithm applied to the circles in the top half, reflecting across $\gamma_1 \cup \gamma_3$. The SAR data for f is close to smooth (b, c), showing the singularities in f nearly cancel. The backprojected image (d) is also much smoother than the original.

BIBLIOGRAPHY

[1] Ambartsoumian G, Felea R, Krishnan V P, Nolan C and Quinto E T 2013 A class of singular Fourier integral operators in radar imaging *J. Funct. Anal.* **264** 246–69

[2] Ambartsoumian G and Kuchment P 2005 On the injectivity of the circular Radon transform *Inverse Problems* **21** 473–85

[3] Agranovsky M L, Kuchment P and Kunyansky L 2009 On reconstruction formulas and algorithms for the thermoacoustic tomography, in *Photoacoustic imaging and spectroscopy* ed Wang L V (Boca Raton: CRC Press) pp 89–101

[4] Agranovsky M L and Quinto E T 1996 Injectivity sets for the Radon transform over circles and complete systems of radial functions *J. Funct. Anal.* **139** 383–414

[5] Agranovsky M L and Quinto E T 2001 Geometry of stationary sets for the wave equation in $\mathbb{R}^n$: the case of finitely supported initial data *Duke Math. J.* **107** 57–84

[6] Bérenger J-P 1994 A perfectly matched layer for the absorption of electromagnetic waves *J. Comput. Phys.* **114** 185–200

[7] Beltukov A and Feldman D 2009 Identities among Euclidean Sonar and Radon transforms *Adv. Appl. Math.* **42** 23–41.

[8] Candès E, Demanet L and Ying L 2009 A fast butterfly algorithm for the computation of Fourier integral operators *Multiscale Model. Simul.* **7** 1727–50

[9] Candès E, Demanet L and Ying L 2007 Fast computation of Fourier integral operators *SIAM J. Sci. Comput.* **29** 2464–93

[10] de Hoop M V, Uhlmann G, Vasy A and Wendt H 2013 Multiscale discrete approximations of Fourier integral operators associated with canonical transformations and caustics *Multiscale Model. Simul.* **11** 566–85

[11] Fawcett J A 1985 Inversion of N-dimensional spherical averages *SIAM J. Appl. Math.* **45** 336–41

[12] Felea R 2007 Displacement of artefacts in inverse scattering *Inverse Problems* **23** 1519–31

[13] Finch D, Haltmeier M and Rakesh 2007 Inversion of spherical means and the wave equation in even dimensions *SIAM J. Math. Anal.* **68** 392–412

[14] Finch D, Patch S K and Rakesh 2004 Determining a function from its mean values over a family of spheres *SIAM J. Math. Anal.* **35** 1213–40

[15] Finch D and Rakesh 2009 Recovering a function from its spherical mean values in two and three dimensions, in *Photoacoustic imaging and spectroscopy* ed Wang L V (Boca Raton: CRC Press) pp 77–88

[16] Grote M J and Sim I 2010 Efficient PML for the wave equation *Preprint* arXiv:1001.0319

[17] Haltmeier M 2014 Exact reconstruction formula for the spherical mean Radon transform on ellipsoids *Preprint* arXiv:1404.3935

[18] Haltmeier M 2014 Universal inversion formulas for recovering a function from spherical means *SIAM J. Math. Anal.* **46** 214–32

[19] Hörmander L 1994 *The analysis of linear partial differential operators* vol IV (*Grundlehren der Mathematischen Wissenschaften* vol 275) corrected reprint of the 1985 original (Berlin: Springer)

[20] Louis A K and Quinto E T 2000 Local tomographic methods in SONAR, in *Surveys on solution methods for inverse problems* ed Colton D *et al* (Vienna: Springer)

[21] Nolan C J and Cheney M 2004 Microlocal analysis of synthetic aperture radar imaging *J. Fourier Anal. Appl.* **10** 133–48

[22] Nolan C J, Cheney M, Dowling T and Gaburro R 2006 Enhanced angular resolution from multiply scattered waves *Inverse Problems* **22** 1817–34.

[23] Nolan C J and Cheney M 2003 Synthetic aperture inversion for arbitrary flight paths and nonflat topography *IEEE Trans. Image Process.* **12** 1035–43

[24] Quinto E T, Rieder A, and Schuster T (2011) Local inversion of the sonar transform regularized by the approximate inverse *Inverse Problems* **27**

[25] Ramm A G 2002 Injectivity of the spherical means operator *C. R. Math. Acad. Sci. Paris* **335** 1033–8

[26] Stefanov P and Uhlmann G 2013 Is a curved flight path in SAR better than a straight one? *SIAM J. Appl. Math.* **73** 1596–612

[27] Yazıcı B, Cheney M and Yarman C E 2006 Synthetic-aperture inversion in the presence of noise and clutter *Inverse Problems* **22** 1705–29

Chapter 3

COMPUTATION OF FOURIER INTEGRAL OPERATORS WITH CAUSTICS

3.1 Introduction

Fourier integral operators have found use in representing solutions of PDEs and in inverse problems. For instance, they model the solution operators of hyperbolic equations, and a variety of integral transforms, such as the Radon transform and its generalizations. With these applications in mind, generic algorithms for computing FIOs would be a useful computational tool for researchers. Many FIOs of interest are associated with a canonical transformation $\chi\colon (x,\xi) \mapsto (y,\eta)$, and these will be the operators considered in this paper. For short, we will call these operators *graph FIOs*.

A graph FIO can be represented locally as an oscillatory integral

$$(Af)(y) = \int_{\mathbb{R}^n} e^{i\phi(y,\xi)} a(y,\xi) \hat{f}(\xi)\, d\xi. \tag{3.1.1}$$

Direct integration of (3.1.1) is possible, but inefficient ($O(N^{2n})$ time), and it ignores the special structure of the operator. Recently, Candès, Demanet, and Ying have given several new algorithms for efficiently computing graph FIOs in dimension 2: an $O(N^{2.5} \log N)$ algorithm [11] based on dividing frequency space into properly-sized wedges, and in [10] an intriguing $O(N^2 \log N)$ algorithm with a butterfly structure.

Smith [31, 32] proved that FIOs associated with canonical transformations map a wave packet concentrated in space and frequency at (x,ξ) to one centered at $\chi(x,\xi)$. This sparsity was noted later also by Candès and Demanet [7, 8], who used curvelets and proved that graph FIOs have essentially sparse representations with respect to the curvelet frame. Later, Guo and Labate [19] showed sparsity for another time- and frequency-localized frame, shearlets,

and Cordero, Gröchenig, and Nicola [13] showed sparsity for a class of tame graph FIOs with respect to a Gabor frame.

These results all suggest that wave packets (and other time- and frequency-localized frames) could provide a useful approach to computing graph FIOs. In [1], Andersson, de Hoop, and Wendt designed a $O(N^{(3n-1)/2} \log N)$ wave packet-based algorithm for computing graph FIOs, building on [14, 15].

These three algorithms have several common elements: each involves splitting the exponential term of 3.1.1 into oscillatory and non-oscillatory parts and constructing a low-rank approximation for the non-oscillatory parts separating the y and ξ variables. Each has a different approach to constructing low-rank approximations: Andersson et al. use prolate spheroidal wave functions (PSWFs). Candès et al. in [11] offer two approaches: a randomized algorithm and a deterministic method using Taylor series. Finally, [10] employs a polynomial approximation using Chebyshev grids.

All three algorithms also require the FIO to be represented in the form (3.1.1). Several issues can occur when representing an FIO this way. In particular, there may be *caustics*, points where Λ (the graph of χ) cannot be locally parameterized by (y, ξ). Even away from caustics, parameterizing the Lagrangian Λ in (y, ξ) coordinates, instead of the natural (x, ξ) coordinates, is messy for a practical implementation.

De Hoop et al. [16] describe how to extend Andersson et al.'s algorithm to overcome caustics, using appropriate changes of coordinates $(x, \xi) \mapsto (\tilde{x}, \tilde{\xi})$. In this paper, we describe a modified version of de Hoop et al.'s algorithm for computing graph FIOs of arbitrary dimension. The key features of the algorithm are its ease of use and automatic handling of caustics. The user of the algorithm, armed only with an FIO's canonical transformation and principal symbol, can immediately compute the operator without performing any calculations or writing extra code. The canonical transformation in many cases is much more natural to work with than a specific phase function. The principal symbol, if unknown, can often be taken as 1 without affecting the singularities of Af. As well, we present companion algorithms for computing the inverse and transpose operators without inverting the canonical

transformation explicitly.

§3.2 gives a mathematical description of the forward algorithm, which is split into local and global components, and then describes how to compute the inverse and transpose. §3.3 applies the algorithm to a variety of FIOs, including the Radon transform, variable-speed wave propagation, and synthetic aperture radar ghost images. §3.4 describes some details of the numerical implementation. §3.5 collects some proofs of statements from the body of the paper. Finally, §3.6 describes potential directions for future research. Our MATLAB implementation, with documentation, is available from the author at **pcaday@math.washington.edu**.

3.2 Mathematical Description

In this section, we describe the algorithm from a mathematician's perspective.

Let X, Y be compact sets in $\mathbb{R}^n$, $n > 2$. Our inputs are a distribution $f \in \mathcal{D}'(X)$ and a graph FIO $A : C^\infty(X) \to \mathcal{D}'(Y)$ of order m. We describe A by its canonical transformation $\chi : T^*X \to T^*Y$ and its principal symbol σ. These of course only characterize A modulo operators of order $m - 1$, but as the algorithm we will develop computes A modulo an FIO of order $m - \frac{1}{2}$, this is not a problem. Furthermore, we assume σ can be taken to be homogeneous of order m in ξ, which covers a wide range of useful FIOs, such as the examples given in section 3.3.

A can locally be described as an oscillatory integral, as in (3.1.1), and for graph FIOs we have an especially convenient local representation, under a certain condition (lack of caustics). We will present an algorithm for approximating such oscillatory integrals in section 3.2.1. To extend this to a global algorithm for computing A, we then describe in 3.2.2 how A can be broken into a sum of integrals of this kind.

Notation. Let $I^m(\chi) = I^m(Y \times X, \chi)$ denote the space of Fourier integral operators of order m with canonical transformation χ: $T^*X \to T^*Y$. Also, throughout $\operatorname{sgn} A$ will denote the signature of A (the sum of the signs of its eigenvalues).

3.2.1 Local Algorithm

Before describing the local algorithm, we recall the (micro)local representation for graph FIOs, where "microlocal" means that it is valid for $\mathrm{WF}(f)$ supported near some $\gamma_0 \in \Lambda'$. Our goal is to write

$$(Af)(y) = \frac{1}{(2\pi)^n} \int e^{iS(y,\xi)} a(y,\xi) \hat{f}(\xi)\, d\xi \tag{3.2.1}$$

We suppose we are given local coordinates x, y on X and Y, inducing coordinates (y, η, x, ξ) on $T^*Y \times T^*X \supset \Lambda'$. Proposition 21.2.18 in Hörmander [20] states that if the projection $\Lambda' \to (y,\xi)$ is a local diffeomorphism, there is a generating function $S(y,\xi)$ (homogeneous in ξ of order 1) such that $S(y,\xi) - x\cdot\xi$ is a phase function for Λ near γ_0. Then, by [20, proposition 25.1.5], if $\mathrm{WF}'(A)$ is supported in a sufficiently small neighborhood of γ_0, there is an amplitude $a(y,\xi)$ so that (3.2.1) holds, modulo C^∞.

The condition for the local algorithm to work, then, is that γ_0 is not a caustic:

Definition 3.1. γ_0 is a *caustic* if the projection $\pi\colon \Lambda' \to (y,\xi)$ fails to be a local diffeomorphism at γ_0. Equivalently, γ_0 is a caustic if and only if $\det \partial y/\partial x \neq 0$.

The equivalence is true because (x,ξ) are global coordinates for Λ'. The good news is that it is always possible to choose local coordinates for X in which γ_0 is not a caustic [20, proposition 25.3.3].

Note that S is uniquely defined by χ. If γ_0 is not a caustic, (y,ξ) are local coordinates for Λ', so we can write $x = x(y,\xi)$ and $\eta = \eta(y,\xi)$. Since the Lagrangian associated with $S - x\cdot\xi$ is $\{y, \frac{\partial S}{\partial y}, \frac{\partial S}{\partial \xi}, -\xi\}$,

$$\frac{\partial S}{\partial \xi} = x(y,\xi), \qquad \frac{\partial S}{\partial y} = \eta(y,\xi). \tag{3.2.2}$$

We will also write $y = y(x,\xi)$ for the y-component of χ, and write $\partial y/\partial x$, $\partial y/\partial \xi$ for its first derivatives.

With these preliminaries done, we now assume γ_0 is not a caustic and A takes the form (3.2.1). Our next task is to approximate this integral.

Cone Decomposition

The idea for the algorithms by Andersson et al. and Candès et al. [1, 11] is to break up our integral (3.2.1) as a sum of integrals over conical regions $\Xi_1, \ldots, \Xi_r$ in frequency space, so that $\xi/|\xi|$ is near a fixed ν_i on each. On each cone, we can split the exponential term of the integrand into an *oscillatory component*, representing a change of coordinates, and a *nonoscillatory component*, representing the application of a pseudodifferential operator. Let's see how this happens.

First, we must create the cones. Choose r approximately uniformly distributed unit covectors $\nu_1, \ldots \nu_r \in \mathbb{S}^{n-1}$, and let $\Xi_i = \{\xi \in \mathbb{R}^n \setminus \{0\} : |\frac{\xi}{|\xi|} - \nu_i| < \delta\}$, for some δ large enough that the Ξ_i cover $\mathbb{R}^n \setminus \{0\}$. We will see later how we should choose r. Then, let $\{\rho_i(\xi)\}$ be a partition of unity subordinate to the Ξ_i, and let $f_i = \rho_i(D)f$ be the result of multiplying f by ρ_i in the frequency domain.

Focusing on a single cone, say cone i, we consider computing Af_i. Since $\hat{f}_i$ is supported on the cone Ξ_i, we can assume $\xi \in \Xi_i$. Choose a rotated coordinate basis $(\xi', \xi''_1, \ldots, \xi''_{n-1})$ in which $\nu_i = (1, 0, \ldots, 0)$. Then, expanding $S(y, \xi) = \xi' S(y, 1, \xi''/\xi')$ in a Taylor series about ν_i and using Euler's theorem gives:

$$S(y,\xi) = \xi \cdot \frac{\partial S}{\partial \xi}(y, \nu_i) + \frac{1}{2}\frac{\xi''^{\top} \cdot \frac{\partial^2 S}{\partial \xi''^2}(y, \nu_i) \cdot \xi''}{\xi'} + O\left(\frac{|\xi''|^3}{|\xi'|^2}\right). \tag{3.2.3}$$

Now, suppose the cones are thin enough that $|\xi''| = O(\sqrt{|\xi'|})$. (This is the *parabolic scaling* condition for wave packets). This is only possible in our numerical setting, where the maximum frequency is limited. If we have an $N \times \cdots \times N$ grid of fixed size, then $|\xi| = O(N)$, and in order to get $|\xi''| = O(\sqrt{|\xi'|})$, we must have $\delta = O(N^{-1/2})$, requiring $r = O(N^{(n-1)/2})$ cones.

With this parabolic scaling assumption, the second and third terms of (3.2.3) are bounded, so these parts of $\exp[iS(y, \xi)]$ do not oscillate. We can group them instead with the amplitude $a(y, \xi)$ in (3.2.1). The first term, however, *is* oscillatory, and it corresponds to a change of coordinates $x \mapsto \frac{\partial S}{\partial \xi}(y, \nu_i)$; for, if we were to remove the other terms and the amplitude

(and integrate over all ξ), we would have

$$\frac{1}{(2\pi)^n}\int e^{i\xi\cdot\frac{\partial S}{\partial\xi}(y,\nu_i)}\hat{f}(\xi)\,d\xi = f\left(\frac{\partial S}{\partial\xi}(y,\nu_i)\right) = f(x(y,\nu_i)). \tag{3.2.4}$$

Now, let us collect all the nonoscillatory terms and the frequency cutoff ρ_i in the symbol b_i:

$$b_i(y,\xi) = a(y,\xi)e^{i(S(y,\xi)-\xi\cdot x(y,\nu_i))}\rho_i(\xi). \tag{3.2.5}$$

Let B_i be the corresponding pseudodifferential operator

$$(B_i f)(x') = \int e^{ix'\cdot\xi}\, b_i(y(x',\nu_i),\xi)\,\hat{f}(\xi)\,d\xi. \tag{3.2.6}$$

Then Af_i can be decomposed into a change of coordinates and the pseudodifferential operator B_i:

$$Af_i(y) = \int e^{iS(y,\xi)}a(y,\xi)\hat{f}_i(\xi)\,d\xi = \int e^{ix(y,\nu_i)\cdot\xi}b_i(y,\xi)\hat{f}(\xi)\,d\xi = (B_i f)(x(y,\nu_i)). \tag{3.2.7}$$

Adding up the contributions from each cone, we have

$$A = \sum_{i=1}^{r} B_i \circ T_{\nu_i}, \qquad \text{where } T_{\nu_i}(y) = x(y,\nu_i). \tag{3.2.8}$$

This is the first step in Candès et al.'s algorithm [11]. Andersson et al.'s algorithm is similar, except that cones are replaced with boxes that divide frequency space both angularly and radially (the *dyadic parabolic decomposition*). These boxes form the basis for Duchkov et al.'s discrete almost-symmetric wave packets [14] and Candès and Donoho's curvelets [12, 9].

Applying the pseudodifferential operator B_i

In equation (3.2.8), computing graph FIOs locally has been reduced to applying coordinate changes and pseudodifferential operators. Numerically computing pseudodifferential operators is a research topic in its own right: see for example [4]. The usual approach to calculating B_i is through low-rank separated representations of $b_i(x',\xi)$, that is, expansions of the form $b_i(x',\xi) \approx \sum_k F_{ik}(y)G_{ik}(\xi)$. These representations are easy to compute numerically with the DFT:

$$b_i(x',D)u \approx \sum_k F_{ik}(x')\,\mathrm{IDFT}[G_{ik}(\xi)\,\mathrm{DFT}[u]]. \tag{3.2.9}$$

Since the DFT assumes periodicity of its input, appropriate zero-padding of u is necessary before applying the DFT. [11] describes two algorithms for computing low-rank separated representations of $b_i(x', \xi)$, one deterministic and one randomized. Andersson et al. use a deterministic approach to generate these separated representations, which we outline now.

Approximating B_i We start by neglecting the third-order and higher terms in (3.2.3). This truncation is equivalent to making an error of order $O(|\xi''|^3/|\xi'|^2) = O(|\xi|^{-1/2})$ in the generating function S. Since $|\exp[ix] - 1| \leq x$, we contribute an error of order $m - \frac{1}{2}$ in ξ to the amplitude $b(x', \xi)$, where m is the order of the FIO. Therefore, by truncating after the second-order terms, we are perturbing A by an FIO which is half an order smoother, and are replacing B_i by the simpler $\Psi\text{DO}\dot{B}_i$ with symbol

$$\dot{b}_i(x', \xi) = \rho_i(\xi) a(y(x', \nu_i), \xi) E_i(x', \xi), \qquad E_i(x', \xi) = \exp\left[\frac{i\xi''^{\top} \cdot \frac{d^2 S}{d\xi''^2}(x', \nu_i) \cdot \xi''}{2\xi'}\right]. \tag{3.2.10}$$

Setting aside the amplitude term $a(y, \xi)$ for now, we are faced with constructing a low-rank separated representation of E_i, namely, functions $F_{ik}(x')$, $G_{ik}(\xi)$ for the approximation

$$E_i(x', \xi) \approx \sum_k F_{ik}(x') G_{ik}(\xi). \tag{3.2.11}$$

Now, the exponent of E_i can be rewritten as a dot product of $\mathbb{R}^{(n-1)^2}$-valued functions in x' and ξ:

$$E_i(x', \xi) = \exp[i\tilde{\mathbf{u}}(x') \cdot \tilde{\mathbf{v}}(\xi)], \qquad \tilde{\mathbf{u}}_{st}(x') = \frac{\partial^2 S}{\partial \xi''_s \, \partial \xi''_t}(x', \nu), \qquad \tilde{\mathbf{v}}_{st}(\xi) = \frac{\xi''_s \xi''_t}{2\xi'} \quad (1 \leq s, t \leq n-1). \tag{3.2.12}$$

Note that, for clarity, the index i is omitted from $\tilde{\mathbf{u}}$ and $\tilde{\mathbf{v}}$.

PSWFs So far, we do not seem to be much closer to our goal of constructing a representation like (3.2.11). However, the form (3.2.12) allows us to use the rich theory of *prolate spheroidal wave functions* (PSWFs) as developed by Slepian, Pollak, and Landau [23, 24, 25,

26]. PSWFs have important application to bandlimited functions (functions with compactly-supported Fourier transform). For us, however, their most important characteristic is that they are the eigenfunctions of the finite Fourier transform operator $\mathcal{E}_c\colon L^2([-1,1]) \to L^2([-1,1])$ with kernel $K_c(x,y) = \exp(icxy)$. As eigenfunctions, they form a complete basis for $L^2([-1,1])$ and allow us to construct exactly the kind of separated representation (3.2.11)) for $\exp(icxy)$ we wanted.

Theorem 3.1 (Slepian et al. [23]). *For all $c > 0$, there is a complete orthonormal basis for $L^2([-1,1])$ of real eigenfunctions $\{\psi_j^{(c)}\}_{j=1}^{\infty}$ of $\mathcal{E}_c$. If $\lambda_j^{(c)}$ are the corresponding eigenvalues, $|\lambda_1^{(c)}| \geq |\lambda_2^{(c)}| \geq \cdots$, then*

$$\exp(icxy) = \sum_{j=0}^{\infty} \lambda_j^{(c)} \psi_j^{(c)}(x) \psi_j^{(c)}(y), \tag{3.2.13}$$

and convergence is uniform.

This gives us a separated representation for $\exp(icxy)$, but for computation we need a good approximate representation, where the infinite sum of (3.2.13) is replaced by a finite one. Fortunately, the λ_j decay very rapidly for large j, while the ψ_j grow slowly. This allows us to achieve a very accurate approximation by truncating the infinite sum (3.2.13). Specifically,

Theorem 3.2 (Landau and Widom [21]). *Let $N_\lambda(c,\epsilon)$ be the number of λ_j greater than ϵ in absolute value. Then*

$$N_\lambda(c,\epsilon) = \frac{2c}{\pi} + \left(\frac{1}{\pi^2}\log\frac{1-\epsilon}{\epsilon}\right)\log c + o(\log c). \tag{3.2.14}$$

Theorem 3.3 (Widom [34]). *For fixed c, as $j \to \infty$*

$$\lambda_j^{(c)} \sim \left(\frac{ec}{4(j+\frac{1}{2})}\right)^{2j+1}. \tag{3.2.15}$$

Theorem 3.4 (Bonami and Karoui [5]). *If $j \geq 2c/\pi$, then*

$$\left\|\psi_j^{(c)}\right\|_\infty \leq \varkappa\sqrt{j+1}, \text{ where } \varkappa = \frac{5}{4}\sqrt{\frac{\pi\sqrt{5}}{2}}. \tag{3.2.16}$$

Theorem 3.2 shows that $\lambda_j^{(c)}$ is near 1 for $j \ll 2c/\pi$, passes through $\frac{1}{2}$ for $j \approx 2c/\pi$, then rapidly approaches zero. Theorem 3.3 shows that the convergence is eventually super-exponential. Hence for an accurate low-rank approximation choosing $j = 2ck/\pi$ for a small constant k is a good choice. Bonami and Karoui [6] have found a quantitative version of theorem 3.3 that gives explicit bounds on $\lambda_j^{(c)}$. Combining theorems 3.3 and 3.4 gives

Corollary 3.5. *For fixed c,*

$$\left|\exp(icxy) - \sum_{j=0}^{J} \lambda_j^{(c)}\psi_j^{(c)}(x)\psi_j^{(c)}(y)\right| = O\left(\frac{ec}{4(J+\frac{1}{2})}\right)^{2J}. \tag{3.2.17}$$

All that remains is to convert (3.2.12) into the proper form. Since the domain X is compact, $\tilde{\mathbf{u}}(x')$ is bounded. As for $\tilde{\mathbf{v}}$, it is bounded on Ξ_i by construction of Ξ_i. (see section 3.2.1). Scaling $\tilde{\mathbf{u}}$ and $\tilde{\mathbf{v}}$ by their largest infinity-norms, we get

$$E_i(x',\xi) = \exp[ic\mathbf{u}(x')\cdot\mathbf{v}(\xi)],$$

$$\mathbf{u}(x') = \frac{\tilde{\mathbf{u}}(x')}{\sup_{x'}\|\tilde{\mathbf{u}}(x')\|_\infty}, \quad \mathbf{v}(\xi) = \frac{\tilde{\mathbf{v}}(\xi)}{\sup_{\xi}\|\tilde{\mathbf{v}}(\xi)\|_\infty}, \quad c = \sup_{x'}\|\tilde{\mathbf{u}}(x')\|_\infty \cdot \sup_{\xi}\|\tilde{\mathbf{v}}(\xi)\|_\infty. \tag{3.2.18}$$

Now, we need the multidimensional version of (3.2.13). Let $\mathbf{j} = (j_1,\ldots,j_N) \in \mathbb{N}^N$ and $\psi_{\mathbf{j}}^{(c)}(x) = \psi_{j_1}^{(c)}(x_1)\cdots\psi_{j_N}^{(c)}(x_N)$. Then, if $x, y \in L^2([-1,1]^N)$,

$$\exp[icx\cdot y] = \prod_{1\leq\ell\leq N}\exp[icx_\ell y_\ell] = \sum_{\mathbf{j}\in\mathbb{N}^N}\lambda_{\mathbf{j}}^{(c)}\psi_{\mathbf{j}}^{(c)}(x)\psi_{\mathbf{j}}^{(c)}(y). \tag{3.2.19}$$

Combining (3.2.19) with (3.2.13) gives

$$E_i(x',\xi) = \sum_{\mathbf{j}\in\mathbb{N}^N}\lambda_{\mathbf{j}}^{(c)}\psi_{\mathbf{j}}^{(c)}(\mathbf{u}(x'))\,\psi_{\mathbf{j}}^{(c)}(\mathbf{v}(\xi)). \tag{3.2.20}$$

Letting $\mathbf{J}(c,\epsilon)$ be the set of $\mathbf{j} = (j_i) \in \mathbb{N}^N$ such that $|j_1 \cdots j_N \lambda_{\mathbf{j}}^{(c)}| \leq \epsilon$, we have the low-rank approximation

$$\left| \exp[icx \cdot y] - \sum_{\mathbf{j} \in \mathbf{J}(c,\epsilon)} \lambda_{\mathbf{j}}^{(c)} \psi_{\mathbf{j}}^{(c)}(x)\, \psi_{\mathbf{j}}^{(c)}(y) \right| = O(\epsilon). \tag{3.2.21}$$

For fixed c, $|\mathbf{J}(c,\epsilon)| = o(\log^N \epsilon)$ in view of Theorems 3.3 and 3.4, while for fixed ϵ, $|\mathbf{J}(c,\epsilon)| = O(c^N)$ in view of Theorem 3.2.

Generalized Prolate Spheroidal Functions Note that [16] (as well as the current implementation of the algorithm) use the unit ball $\overline{\mathbb{B}}^N$ instead of the unit hypercube $[-1,1]^N$ as the domain for $\tilde{\mathbf{u}}$, $\tilde{\mathbf{v}}$. This requires a new basis of eigenfunctions, those for the truncated Fourier transform on $L^2(\overline{\mathbb{B}}^N)$, which are Slepian's *generalized prolate spheroidal functions* [26]. However, the rest of the algorithm is completely independent of the choice of domain for $\tilde{\mathbf{u}}$ and $\tilde{\mathbf{v}}$. The unit hypercube is advantageous theoretically for establishing bounds on the eigenfunctions and eigenvalues of the truncated Fourier transform, since PSWFs have been better studied than generalized PSWFs.

Amplitude Now that we have an effective algorithm for computing the $E_i(x',\xi)$ component of the ΨDO, we turn to the $a(y(x',\nu),\xi)$ term. Again, we can do a Taylor expansion around the center of the cone. Given ξ, we let $\xi_0 = \nu|\xi|$; then

$$a(y,\xi) = a(y,\xi_0) + (\xi - \xi_0) \cdot \frac{\partial a}{\partial \xi}(y,\xi_0) + (\xi - \xi_0)^\top \frac{\partial^2 a}{\partial \xi^2}(y,\xi^*(y))(\xi - \xi_0) \tag{3.2.22}$$

for some $\xi^*(y)$ on the line segment between ξ_0 and ξ. The most basic possible approximation is the zeroth-order approximation

$$a(y,\xi) \approx a(y,\xi_0) = |\xi|^m a(y,\nu_i). \tag{3.2.23}$$

(Since the principal symbol is homogeneous by assumption, so is a.) Because $|\xi - \xi_0| = O(|\xi|^{1/2})$ and $\frac{\partial a}{\partial \xi} = O(|\xi|^{m-1})$, this approximation has $O(|\xi|^{m-1/2})$ error, and therefore perturbs A by an FIO of order one-half lower. This is the same error we made with the

approximation of $S(y,\xi)$ earlier, so we consider this an acceptable approximation. Hence the actual ΨDO we will approximate, $\ddot{B}_i$, is the one with symbol

$$\ddot{b}_i(x',\xi) = \rho_i(\xi)\ |\xi|^{\,m}\, a(y(x',\nu_i),\nu_i)\, E_i(x',\xi). \tag{3.2.24}$$

Computation of $\ddot{B}_i$ is via the low-rank separated approximation $\ddot{B}_{i,\epsilon}$ with symbol

$$\ddot{b}_{i,\epsilon}(x',\xi) = \rho_i(\xi)\ |\xi|^{\,m}\, a(y(x',\nu_i),\nu_i) \sum_{\mathbf{j}\in\mathbf{J}(c,\epsilon)} \lambda_{\mathbf{j}}^{(c)} \psi_{\mathbf{j}}^{(c)}(\mathbf{u}(x'))\, \psi_{\mathbf{j}}^{(c)}(\mathbf{v}(\xi)). \tag{3.2.25}$$

Before proceeding further we make a few notes on $\tilde{\mathbf{u}}$ and $\tilde{\mathbf{v}}$.

Notes.

1. Since $\partial S/\partial\xi = x(y,\xi)$ by (3.2.2) we can write

$$\tilde{\mathbf{u}}_{st}(x') = \nu_s''^{\top} \frac{\partial x(y,\xi)}{\partial \xi} \nu_t'', \tag{3.2.26}$$

   where ν_s'' is the unit vector in the ξ_s'' direction. We can write $\tilde{\mathbf{u}}$ more directly in terms of the canonical map $\chi = (y(x,\xi),\eta(x,\xi))$ as:

$$\tilde{\mathbf{u}}_{st} = -\nu_s''^{\top} \left(\frac{\partial y(x,\xi)}{\partial x}\right)^{-1} \left(\frac{\partial y(x,\xi)}{\partial \xi}\right) \nu_t'' \Bigg|_{x=x(y,\xi).} \tag{3.2.27}$$

2. Since $\tilde{\mathbf{u}}_{st} = \tilde{\mathbf{u}}_{ts}$ and $\tilde{\mathbf{v}}_{st} = \tilde{\mathbf{v}}_{ts}$, we can restrict to $s \le t$ in (3.2.12), and multiply one of them (say $\tilde{\mathbf{v}}_{st}$) by 2 for $s \ne t$. This has the advantage of reducing the dimension of $\tilde{\mathbf{u}}$, $\tilde{\mathbf{v}}$ to $\binom{n-1}{2}$, decreasing the number of terms needed in the low-rank approximation (3.2.21).

3. The supremum in the definition of $\mathbf{u}$ in (3.2.18) can be taken over the x-support of $a(y(x,\nu),\nu)$ instead of all X, in view of the amplitude approximation (3.2.23).

Local approximation formula

Now that we have reasonable approximations (3.2.20), (3.2.23) for the exponential and amplitude parts of B_i, we can substitute them into equation (3.2.8) for the original FIO A.

Recalling that $\hat{f}_i = \rho_i \hat{f}$ and attaching i indices to c, $\mathbf{u}$, and $\mathbf{v}$, we have

$$Af(y) \approx \sum_{i=1}^{r} \sum_{\mathbf{j} \in \mathbf{J}(c,\epsilon)} \lambda_{\mathbf{j}}^{(c_i)} \psi_{\mathbf{j}}^{(c_i)}(\mathbf{u}_i(x')) \, \mathcal{F}^{-1}_{\xi \to x'} \Big[\psi_{\mathbf{j}}^{(c_i)}(\mathbf{v}_i(\xi)) \, |\xi|^m \rho_i(\xi) \hat{f}(\xi) \Big] a(y, \nu_i) \Big|_{x' = x(y,\nu_i).} \tag{3.2.28}$$

Here $\mathcal{F}^{-1}_{\xi \to x'}$ represents the inverse Fourier transform.

Now the local approximation formula (3.2.28) is independent of γ_0, depending only on χ and a, which can be deduced from the principal symbol. Hence, while (3.2.1) was originally valid only for $\mathrm{WF}'(A)$ in a small neighborhood of γ_0, by a partition of unity the local approximation formula is valid whenever $\mathrm{WF}'(A)$ lies entirely in an open subset $U \subseteq \Lambda'$ that can be parameterized by (y, ξ).

We can go further and relax the assumption that U can be parameterized by (y, ξ). First, replace the pullback in (3.2.28) with a pushforward (in the sense of distributions) by the inverse map, $x \mapsto y(x, \nu_i)$ and a Jacobian factor. Using the pushforward is also advantageous practically, as it allows us to avoid inverting the canonical transformation. Then, dropping the prime on x',

$$Af(y) \approx \mathcal{A}_\epsilon f(y) = \sum_{i=1}^{r} \Big(y(\cdot, \nu_i) \Big)_* \left| \det \frac{\partial y}{\partial x}(x, \nu_i) \right| \cdot \sum_{\mathbf{j} \in \mathbf{J}(c_i,\epsilon)} \lambda_j^{c_i} \psi_j^{c_i}(\mathbf{u}_i(x)) \, \mathcal{F}^{-1}_{\xi \to x} \Big[\psi_j^{c_i}(\mathbf{v}_i(\xi)) \, |\xi|^m \rho_i(\xi) \hat{f}(\xi) \Big] a(x, \nu_i). \tag{3.2.29}$$

Note that we have abused notation and written $a(x, \nu_i) = a(y(x, \nu_i), \nu_i)$. This makes sense if we think of a in connection with the principal symbol, which is defined on the Lagrangian Λ', for which (x, ξ) are global coordinates.

Next, drop the assumption that (y, ξ) parameterizes U and only assume U contains no caustics; in other words, $\det \partial y / \partial x \neq 0$. Note that equation (3.2.29) still makes sense; it only depends on Λ via $\mathbf{u}$, which by (3.2.27) can be defined directly in terms of the canonical transformation χ.

Now, $\det \partial y / \partial x \neq 0$ implies that U can locally be parameterized by (y, ξ). Covering U by open sets that can be parameterized by (y, ξ) and taking a partition of unity, we find that (3.2.29) is valid for all of U. Specifically, if $1 = \sum \alpha_i$ is the partition of unity and U_i are

the open sets, approximation (3.2.29) is valid with a replaced by $a\alpha_i$ for each i. Summing over i and using linearity implies the formula is valid for a also.

Summarizing, we have developed the following approximation.

Theorem 3.6 (Local approximation algorithm)**.** *Let $A \in I^m(Y \times X)$ be an FIO associated with the canonical transformation $\chi\colon T^*X \to T^*Y$, such that $\det \partial y/\partial x \neq 0$ on $\mathrm{WF}'(A)$.*

There exists an error operator $K \in I^{m-1/2}(Y \times X, \chi)$ such that for small $\epsilon > 0$,

$$\frac{\|Af - Kf - \mathcal{A}_\epsilon f\|_2}{\|f\|_{H^m}} = O(\epsilon), \tag{3.2.30}$$

where $\mathcal{A}_\epsilon f$ is the approximation defined in (3.2.29).

The error operator K, of course, corresponds to the higher-order terms in the Taylor expansions of amplitude and phase that were neglected.

Proof. With $\mathcal{A}_\epsilon$ defined as in (3.2.29), and recalling (3.2.8), we have

$$Af(y) = \sum_{i=1}^{r} T_{\nu_i} \circ B_i, \tag{3.2.31}$$

$$\mathcal{A}_0 f(y) = \sum_{i=1}^{r} T_{\nu_i} \circ \ddot{B}_i, \tag{3.2.32}$$

$$\mathcal{A}_\epsilon f(y) = \sum_{i=1}^{r} T_{\nu_i} \circ \ddot{B}_{i,\epsilon}. \tag{3.2.33}$$

As we noted earlier, the symbols b_i and $\ddot{b}_i$ of B_i and $\ddot{B}_i$, respectively, differ by a symbol of order $m - \frac{1}{2}$, and $K = A - \mathcal{A}_0$ is an FIO of order $m - \frac{1}{2}$. All that remains for us is to estimate $\mathcal{A}_0 - \mathcal{A}_\epsilon$.

First, we estimate $\ddot{B}_i - \ddot{B}_{i,\epsilon}$:

$$(\ddot{B}_i - \ddot{B}_{i,\epsilon})f(x') = \sum_{\mathbf{j} \notin \mathbf{J}(c_i,\epsilon)} \lambda_{\mathbf{j}}^{(c_i)} \psi_{\mathbf{j}}^{(c_i)}(\mathbf{u}(x')) \mathcal{F}^{-1}\left[\psi_{\mathbf{j}}^{(c_i)}(\mathbf{v}(\xi))\rho_i(\xi)\,|\xi|^{\,m}\hat{f}(\xi)\right]. \tag{3.2.34}$$

Since

$$\begin{aligned}\left\|\mathcal{F}^{-1}\left[\psi_{\mathbf{j}}^{(c_i)}(\mathbf{v}(\xi))\rho_i(\xi)\,|\xi|^{\,m}\hat{f}(\xi)\right]\right\|_2 &= \left\|\psi_{\mathbf{j}}^{(c_i)}(\mathbf{v}(\xi))\rho_i(\xi)\,|\xi|^{\,m}\hat{f}(\xi)\right\|_2 \\ &\leq \left\|\psi_{\mathbf{j}}^{(c_i)}\right\|_\infty \|\rho_i\|_\infty \left\||\xi|^{\,m}\hat{f}(\xi)\right\|_2 \\ &\leq \left\|\psi_{\mathbf{j}}^{(c_i)}\right\|_\infty \|f\|_{H^m}.\end{aligned} \tag{3.2.35}$$

we have

$$\left\|\lambda_{\mathbf{j}}^{(c_i)}\psi_{\mathbf{j}}^{(c_i)}(\mathbf{u}(x'))\mathcal{F}^{-1}\left[\psi_{\mathbf{j}}^{(c_i)}(\mathbf{v}(\xi))\rho_i(\xi)\,|\xi|^{\,m}\hat{f}(\xi).\right]\right\|_2 \leq \left|\lambda_{\mathbf{j}}^{(c_i)}\right|\left\|\psi_{\mathbf{j}}^{(c_i)}\right\|_\infty^2 \|f\|_{H^m}. \tag{3.2.36}$$

□

3.2.2 Global Algorithm

In the previous section, we derived a local approximation (3.2.29) for A, which is valid as long as $\mathrm{WF}'(A)$ contains no caustics. Our goal in this section is to derive a global approximation algorithm for A by expressing A as a sum of oscillatory integrals without caustics to which (3.2.29) applies. To handle caustics, de Hoop et al. [16] use appropriate changes of coordinates in X, which they call *singularity-resolving diffeomorphisms*. Note that their approach could also be combined with Candès et al.'s wedge algorithm [11] and potentially their butterfly algorithm [10].

Throughout, let x denote a fixed choice of coordinates on X, and $\tilde{x}$ another choice of coordinates. Let (x,ξ) and $(\tilde{x},\tilde{\xi})$ denote the associated coordinates on T^*X.

First Draft

By proposition 25.3.3 in Hörmander [20], for each $\gamma_0 \in \Lambda'$ there is always a local choice of coordinates $\tilde{x}$ for X in which Λ' does not have caustics in a neighborhood of γ_0. This gives us the following first draft of an algorithm for computing Af with caustics present.

For each $\gamma \in \Lambda'$ we get a neighborhood U_γ with associated caustic-free coordinates $\tilde{x}^{(\gamma)}$. We cover Λ' by the U_i, which we can assume to be conic (since $\partial y/\partial x$ is homogeneous in ξ). Since X, Y are compact, the cosphere bundle $S^*X \times S^*Y$ is also compact, so $\Lambda' \cap (S^*X \times S^*Y)$ is compact, and we can choose a finite subcover $U_{i_1} \cap (S^*X \times S^*Y), \ldots, U_{i_r} \cap (S^*X \times S^*Y)$. Then $U_{i_1}, \ldots, U_{i_r}$ covers Λ; choosing a partition of unity ρ_j subordinate to this subcover, we can write $Af = A_1 f + \cdots + A_r f$, where A_j is an FIO with canonical transformation χ associated with the principal symbol $\rho_j\sigma$. (Recall σ is the principal symbol of A.) Each $A_j f$ can be computed by theorem 3.6.

Quadratic Coordinate Changes

While the first draft algorithm is valid, it does not specify how to find caustic-free coordinates. However, this turns out to be an easy task: we can identify a specific finite family of simple coordinate changes that are sufficient for all A. If $\varphi : \mathbb{R}^n \to \mathbb{R}^n$ is a diffeomorphism, and $\tilde{x} = \varphi(x)$, then

$$\frac{\partial y}{\partial \tilde{x}} = \frac{\partial y}{\partial x}(d\varphi)^{-1} + \sum_{k=1}^{n} \frac{\partial y}{\partial \xi_k} \frac{\partial (d\varphi)^\top}{\partial x^k} \xi. \tag{3.2.37}$$

The simplest possible useful diffeomorphisms are therefore those with quadratic terms:

Definition 3.2. Let $t \in \mathbb{R} \setminus \{0\}$ be a fixed constant and $J \subsetneq \{1, \dots, n\}$, $i \in \{1, \dots, n\} \setminus J$. Define the diffeomorphisms $\beta_{i,J} \colon \mathbb{R}^n \to \mathbb{R}^n$ as follows:

$$\beta_{i,J}(x_1, \dots, x_n) = \left(x_1, \dots, x_{i-1}, x_i - \frac{t}{2} \sum_{j \in J} x_j^2, x_{i+1}, \dots, x_n \right). \tag{3.2.38}$$

It turns out that in all cases, at least one of these quadratic coordinate changes avoids a caustic.

Proposition 3.7. *For all $\gamma_0 \in \Lambda'$, there exist $J \subsetneq \{1, \dots, n\}$ and $i \in \{1, \dots, n\} \setminus J$ such that $\det \partial y / \partial \tilde{x}^{(i,J)}(\gamma_0) \neq 0$, where $\tilde{x}^{(i,J)} = \beta_{i,J}(x)$.*

Remark. There are $1 + n(2^{n-1} - 1)$ of these diffeomorphisms; so 3 in dimension 2 and 10 in dimension 3.

A proof is given in §3.5. Note that while proposition 3.7 guarantees that the $\beta_{i,J}$ always suffice for avoiding caustics, the choice of coordinates will affect the quality of the algorithm's output. With this in mind, we will from here on consider a generic set of coordinates $\tilde{x}^{(1)}, \dots, \tilde{x}^{(Q)}$ such that no point on γ is simultaneously a caustic with respect to all $\tilde{x}^{(j)}$. Proposition 3.7 gives us one such coordinate set.

With the coordinate set chosen, define $U_k = \{(x, \xi) \in T^*X : \det \partial y / \partial \tilde{x}^{(k)} \neq 0\}$, and choose a partition of unity $\{\alpha_k\}_{k=1}^{Q}$ subordinate to the U_k. As before, A can be written as $A_1 + \cdots + A_r$, where A_k is an FIO associated with the canonical transformation χ, with principal symbol $\alpha_k \sigma$. Each $A_k f$ can then be approximated via theorem 3.6.

Although it is only necessary that $\alpha_k = 0$ when $\det \partial y/\partial \tilde{x}^{(k)} = 0$, it is advantageous numerically for α_k to be zero when $\partial y/\partial \tilde{x}^{(k)}$ is close to singular. Equation (3.2.27) shows that $\|\tilde{\mathbf{u}}\|$ can be large when $\partial y/\partial \tilde{x}^{(k)}$ is nearly singular. Large $\tilde{\mathbf{u}}$, in turn, increase c (eq. (3.2.18)), which increases the number of significant terms in the PSWF decompositions (eq. (3.2.15)). By applying more restrictive cutoffs α_k, we can potentially avoid cons Ξ_i where c may be large, reducing computation time.

One way to choose the α_k to achieve this goal is as follows: compute the minimum singular values $\mu_k(x,\xi)$ of $\partial y/\partial \tilde{x}^{(k)}$, and let $\mu(x,\xi) = \max_k \mu_k(x,\xi)$. Then, choose $\tilde{\alpha}_k \in C^\infty$ such that

$$\tilde{\alpha}_k(x,\xi) = \begin{cases} 0, & \frac{\mu_k}{\mu} < 1 - \tau_w, \\ 1, & \frac{\mu_k}{\mu} > 1 - \frac{1}{10}\tau_w, \end{cases} \tag{3.2.39}$$

where $\tau_w \in (0,1)$ is a parameter controlling the tightness of the partition. By forcing $\alpha_k = 1$ when $\frac{\mu_k}{\mu}$ lies in a neighborhood of 1, we are assured that $\tilde{\alpha}_k$ is smooth. Finally, define $\alpha_k = \tilde{\alpha}_k/(\tilde{\alpha}_1 + \cdots + \tilde{\alpha}_Q)$ to complete the partition of unity construction.

Amplitude Calculation

The last major ingredient for the algorithm we have neglected to discuss is how to construct the amplitude a for (3.2.1) from the principal symbol. This is standard FIO theory, but using the principal symbol in a numerical algorithm requires consideration. The principal symbol is defined as a symbol on the half-density bundle on $T^*(X \times Y)$ tensored with the Maslov line bundle. In order to use it in a practical algorithm we will need to trivialize these two bundles in some way.

For graph FIOs, the half-density bundle has a natural global trivialization [20, following (25.3.2)′] in which the pullback of $|dy|^{1/2}|d\eta|^{1/2}$ by the projection $\pi_Y \colon \Lambda \to T^*Y$ is taken as the unit section. This half-density is invariant under changes of coordinates, since the Jacobian of any symplectomorphism $(y,\eta) \mapsto (y',\eta')$ has determinant 1. Equally, we could have used $|dx|^{1/2}|d\xi|^{1/2}$: since Λ is a Lagrangian submanifold of $T^*(X \times Y)$ with respect to the sum of the symplectic forms on T^*X and T^*Y, it follows that $\iota^*\sigma_{T^*X} = -\iota^*\sigma_{T^*Y}$ on Λ.

But the volume forms on T^*X and T^*Y are $\sigma^n_{T^*X}$ and $\sigma^n_{T^*Y}$, so after pulling them back to Λ, the two volume densities are identical: $|dx||d\xi| = |dy||d\eta|$.

It is also possible to globally trivialize the Maslov bundle [18, lemma 4.1.3], but there is not a natural choice of unit section in general. Several special cases do have natural unit sections, such as when the projection of Λ onto $X \times Y$ has constant rank [18, p. 96], which for us is equivalent to $\partial y/\partial \xi$ having constant rank. Other cases having natural global trivializations are when χ is homotopic to the identity [28, ex. 3.7] or when $\mathrm{WF}'(A) \subseteq T^*X \setminus 0$ is simply-connected [28, p.284]. One example of the latter case is the inversion of the circular Radon transform, where $\mathrm{WF}'(A) = \{(r, s, \rho, \sigma) \mid |\sigma| \leq |\rho|,\ r > 0\}$ is the disjoint union of two simply-connected regions.

If we allow for discontinuous sections of the Maslov bundle, however, natural global trivializations are possible. Following Hörmander [20, before (25.1.9)], suppose two local representations of a Fourier integral distribution u in local coordinates $x^{(1)}$, $x^{(2)} \in \mathbb{R}^n$ are given,

$$u(x^{(1)}) = (2\pi)^{-(n+2N_1)/4} \int_{\mathbb{R}^{N_1}} e^{i\phi_1(x^{(1)},\theta_1)} a_1(x^{(1)}, \theta_1)\, d\theta_1, \tag{3.2.40}$$

$$u(x^{(2)}) = (2\pi)^{-(n+2N_2)/4} \int_{\mathbb{R}^{N_2}} e^{i\phi_2(x^{(2)},\theta_2)} a_2(x^{(2)}, \theta_2)\, d\theta_2. \tag{3.2.41}$$

If u has order m, Hörmander shows

$$a_2 d_{C_2}^{1/2} - e^{\pi i s/4} a_1 d_{C_1}^{1/2} \in S^{m-1}(\Lambda, \Omega_\Lambda^{1/2}), \tag{3.2.42}$$

where

$$s = \operatorname{sgn} d^2_{\theta_1} \phi_1(x^{(1)}, \theta_1) - \operatorname{sgn} d^2_{\theta_2} \phi_2(x^{(2)}, \theta_2), \tag{3.2.43}$$

and d_{C_i} is the quotient density on $C_i = \ker d_{\theta_i}\phi_i$ defined by $d_{\theta_i}\phi_i$ with respect to the Euclidean volume densities on $\mathbb{R}^d \times \mathbb{R}^{N_i}$. Then, C_i can be identified with Λ via the isomorphism $T_i \colon C_i \ni (x, \theta_i) \mapsto (x, d_x\phi(x, \theta_i)) \in \Lambda$. Similarly, the signature terms of (3.2.43) can be interpreted as functions on Λ by this isomorphism.

Recalling our choice of unit section for the half-density bundle, $|dx|^{1/2}|d\xi|^{1/2}$, and dividing by it, we can get an expression for the principal symbol from (3.2.42):

$$\sigma = a_i \frac{d_{C_i}^{1/2}}{|dx|^{1/2}|d\xi|^{1/2}} \exp\left[\frac{i\pi}{4} \operatorname{sgn} d^2_{\theta_i}\phi_i\right] \in S^m/S^{m-1}. \tag{3.2.44}$$

In our situation, we have two local forms for graph FIOs in different coordinates. Here the θ_i are the $\xi^{(i)}$ corresponding to different coordinates $x^{(i)}$, and $\phi_i(x, y, \xi) = S_i(y, \xi) - x \cdot \xi$. For graph FIOs, Hörmander shows

$$d_{C_i} = |dy|\,|d\xi| = \left|\det \frac{\partial y}{\partial x}\right| |dx|\,|d\xi|\,.$$

Now, since we have a nice unit section $|dx|^{1/2}|d\xi|^{1/2}$ for Λ, we can remove it, leaving

$$\sigma = a_i \left|\det \frac{\partial y}{\partial x^{(i)}}\right|^{1/2} \exp\left[\frac{i\pi}{4} \operatorname{sgn} d^2_{\xi_i}\phi_i\right],$$

Finally, since $d_\xi \phi = d_\xi S - x = x(y, \xi) - x$, we have $d^2_\xi \phi = \frac{\partial x}{\partial \xi} = -(\frac{\partial y}{\partial x})^{-1}(\frac{\partial y}{\partial \xi})$. Dropping the subscripts,

$$\sigma = a \left|\det \frac{\partial y}{\partial x}\right|^{1/2} \exp\left[-\frac{i\pi}{4} \operatorname{sgn}\left(\frac{\partial y}{\partial x}\right)^{-1}\left(\frac{\partial y}{\partial \xi}\right)\right], \tag{3.2.45}$$

Note that σ is discontinuous unless $\partial y/\partial \xi$ has constant rank. From a numerical standpoint, this would indicate that care should be taken where $\partial y/\partial \xi$ drops rank. In practice, this discontinuous trivialization appears to work satisfactorily. For pseudodifferential operators, σ agrees with the usual principal symbol for pseudodifferential operators, pulled back from T^*X to Λ.

Safarov develops a slightly different discontinuous trivialization, the *singular principal symbol*, in [28], and proves formulas for the singular principal symbols of products and adjoints of graph FIOs. The current implementation of the algorithm supports both trivializations.

3.2.3 Final Algorithm

Combining everything so far, we can present the final approximation:

$$Af(y) \approx \mathcal{A}_\epsilon f(y) = \sum_{k=1}^{Q}\sum_{i=1}^{r_k}\left(y\circ\beta_k^{-1}(\cdot,\tilde{\nu}_{i,k})\right)_*\Bigg[\left|\det\frac{\partial y}{\partial\tilde{x}^{(k)}}(\tilde{x}^{(k)},\tilde{\nu}_{i,k})\right|\cdot\sum_{\mathbf{j}\in\mathbf{J}(c_{i,k},\epsilon)}\lambda_{\mathbf{j}}^{c_{i,k}}\psi_{\mathbf{j}}^{c_{i,k}}(\mathbf{u}_{i,k}(\tilde{x}))\cdot$$
$$\mathcal{F}^{-1}_{\tilde{\xi}\to\tilde{x}}\left[\psi_{\mathbf{j}}^{c_{i,k}}(\mathbf{v}_{i,k}(\tilde{\xi}))\cdot|\tilde{\xi}|^m\rho_{i,k}(\tilde{\xi})\widehat{\beta_k^*f}(\tilde{\xi})\right]\cdot(a_k\circ\beta_k^{-1})(\tilde{x}^{(k)},\tilde{\nu}_{i,k})\Bigg], \quad (3.2.46)$$

where

$$a_k = \alpha_k\sigma_k\left(\det\frac{\partial y}{\partial x}\right)^{-1/2}\exp\left[\frac{i\pi}{4}\operatorname{sgn}\left(\frac{\partial y}{\partial x}\right)^{-1}\left(\frac{\partial y}{\partial\xi}\right)\right]. \quad (3.2.47)$$

Here k indexes the choices of coordinates, i the cones, and $\mathbf{j}$ the PSWFs. As a more readable form, the pseudocode listing 2 indicates how the global approximation formula translates to an algorithm.

Theorem 3.8. *Suppose the values f are given on an $L\times\cdots\times L$ grid in X, and the values of Af are to be calculated on an $M\times\cdots\times M$ grid in Y. Assume the bandwidths $c_{i,k}$, the PSWFs $\psi_{\mathbf{j}}^{(c_{i,k})}$, and eigenvalues $\lambda_{\mathbf{j}}^{(c_{i,k})}$ are precomputed.*

Computing $\mathcal{A}_\epsilon f$ using approximation (3.2.46) and DFTs requires worst-case time

$$O(L^{(3n-1)/2}\log L\log^{n(n-1)/2}\epsilon + L^{(n-1)/2}M^n). \quad (3.2.48)$$

Proof. After the precomputation steps, the global algorithm reduces to applying the local algorithm $Q = O(1)$ times. Assume each pullback β_k^*f is computed on a grid of size $O(L)$. In the local algorithm, the bulk of the computational work comes from computing the PSWF contributions and the pushforwards. For each of the $O(L^{(n-1)/2})$ cones, there are $\mathcal{J}(c,\epsilon) = O(\log^{n(n-1)/2})$ significant PSWF terms, each requiring $O(L^n\log L)$ work (pointwise multiplications and an inverse DFT). The pushforward step in each cone requires $O(L^n + \kappa M^n)$ time using the algorithm described in section 3.4.2, where κ is a constant depending only on the FIO: κ is the maximum number of preimages of a point in Y under $\tilde{x}^{(i)}\mapsto y(\tilde{x}^{(i)},\tilde{\xi}^{(i)}$ for each i and all $\tilde{\xi}^{(i)}$. □

Pseudocode 2 Approximate Af

Precompute α_k, a_k, $c_{i,k}$, $\mathbf{u}_{i,k}$, and PSWFs.
for each coordinate change $k = 1$ **to** Q **do**
  Pull back f by β_k.
  Zero pad $\beta_k^* f$, and apply DFT.
  for each cone $i = 1$ **to** r_k **do**
    for each significant PSWF $j \in \mathbf{J}_{c,e}$ **do**
      Multiply $\widehat{\beta_k^* f}$ by $\rho_{i,k}|\tilde{\xi}|^m$ and the PSWF $\psi_j^{c_{i,k}}(\mathbf{v}_{i,k}(\tilde{\xi}))$.
      Apply inverse DFT.
      Multiply by $\lambda_j^{c_{i,k}} \psi_j^{c_{i,k}}(\mathbf{u}_{i,k})$.
      Add to current sum.
    end for
    Apply pushforward (with Jacobian factor) by $y \circ \beta_k^{-1}$.
    Add to current sum.
  end for
end for
Add to current sum.

3.2.4 Inverse and Transpose

In this section, we develop algorithms parallel to the previous section for computing $A^\top g$ and $A^{-1} g$ (assuming A is elliptic) given the same information as before: the canonical transformation χ and the principal symbol σ of A. Theoretically these algorithms are superfluous, of course, since A^{-1} and $A^\top$ are graph FIOs associated with the inverse canonical transformation χ^{-1} (While χ may not be one-to-one, $\det \chi = 1$ since χ is a symplectomorphism, so χ is locally invertible, and $A^\top$, A^{-1} can be represented as locally finite sums of graph FIOs).

However in practice, computing χ^{-1} may be difficult or inefficient, and if χ^{-1} is not a graph, it must be decomposed as a union of graphs. Moreover, these algorithms come essen-

tially for free: A^{-1} and $A^\top$ can be computed with the same information and precomputation steps used previously to compute A. Throughout A is as before, and $g = g(y)$ an arbitrary distribution on Y.

Transpose

Recall that the overall structure of the local algorithm for Af involved writing A as a sum of ΨDO's with coordinate changes. Originally, from (3.2.8) we have $Af(y) = \sum_{i=1}^r (B_i f_i)(x(y,\nu_i))$. Later, we approximated the ΨDO B_i by $\ddot{B}_i$, which had a convenient separated representation, and interpreted $x(y,\nu_i)$ as a pushforward by $x \mapsto y(x,\nu_i)$ (along with a Jacobian factor). Hence the local forward algorithm can be expressed as:

$$Af(y) \approx \sum_{i=1}^r y(\cdot,\nu_i)_* \left[(\ddot{B}_i f_i) \cdot \det \frac{\partial y}{\partial x} \right]. \tag{3.2.49}$$

The global algorithm is similar, but with coordinate changes $\beta_k \colon x \mapsto \tilde{x}^{(k)}$, and a microlocal partition of unity $\alpha_k \colon T^*X \to [0,1]$. We can express it as

$$Af(y) \approx \sum_{k=1}^Q \sum_{i=1}^{r_k} y \circ \beta_k^{-1}(\cdot,\tilde{\nu}_{i,k})_* \left[\ddot{B}_{i,k}\Big((\beta_k^{-1})^* f \Big) \cdot \det \frac{\partial y}{\partial \tilde{x}^{(k)}} \right]. \tag{3.2.50}$$

where $\ddot{B}_{i,k}$ is the same as $\ddot{B}_i$ in $(\tilde{x}^{(k)}, \tilde{\xi}^{(k)})$ coordinates, except that the symbol is multiplied by the microlocal partition of unity, α_k. Now we can take the transpose; if $f \in C_0^\infty(X)$, $g \in C_0^\infty(Y)$, then

$$\langle Af, g \rangle \approx \sum_{k=1}^Q \sum_{i=1}^{r_k} \left\langle y \circ \beta_k^{-1}(\cdot,\tilde{\nu}_{i,k})_* \left[\ddot{B}_{i,k}\Big((\beta_k^{-1})^* f \Big) \cdot \det \frac{\partial y}{\partial \tilde{x}^{(k)}} \right], g \right\rangle \tag{3.2.51}$$

$$= \sum_{k=1}^Q \sum_{i=1}^{r_k} \left\langle \ddot{B}_{i,k}\Big((\beta_k^{-1})^* f \Big) \cdot \det \frac{\partial y}{\partial \tilde{x}^{(k)}}, y \circ \beta_k^{-1}(\cdot,\tilde{\nu}_{i,k})^* g \right\rangle \tag{3.2.52}$$

$$= \sum_{k=1}^Q \sum_{i=1}^{r_k} \left\langle f, (\beta_k^{-1})_* \ddot{B}_{i,k}^\top \left(\det \frac{\partial y}{\partial \tilde{x}^{(k)}} \right) y \circ \beta_k^{-1}(\cdot,\tilde{\nu}_{i,k})^* g \right\rangle, \tag{3.2.53}$$

so

$$A^\top g \approx \sum_{k=1}^Q \sum_{i=1}^{r_k} (\beta_k^{-1})_* \ddot{B}_{i,k}^\top \left(\det \frac{\partial y}{\partial \tilde{x}^{(k)}} \right) y \circ \beta_k^{-1}(\cdot,\tilde{\nu}_{i,k})^* g. \tag{3.2.54}$$

Recalling the separated representation of the symbol $\ddot{b}_{i,k}$ (the superscripts on $\tilde{x}$ and $\tilde{\xi}$ are omitted for clarity),

$$\ddot{b}_{i,k} = \sum_{\mathbf{j}\in\mathbb{N}^N} (\rho_{i,k}a_k)(\tilde{x}, \tilde{\nu}_{i,k})\lambda_{\mathbf{j}}\psi_{\mathbf{j}}^{c_{i,k}}(\mathbf{u}(\tilde{x}))\psi_{\mathbf{j}}^{c_{i,k}}(\mathbf{v}(\tilde{\xi})), \tag{3.2.55}$$

we have

$$\ddot{B}_{i,k} = \sum_{\mathbf{j}\in\mathbb{N}^N} (\rho_{i,k}a_k)(\tilde{x}, \tilde{\nu}_{i,k})\lambda_{\mathbf{j}}\psi_{\mathbf{j}}^{c_{i,k}}(\mathbf{u}(\tilde{x}^{(k)})) \circ \mathcal{F}^{-1} \circ \psi_{\mathbf{j}}^{c_{i,k}}(\mathbf{v}(\tilde{\xi}^{(k)})) \circ \mathcal{F} \tag{3.2.56}$$

where the first and third operators in the composition are multiplication operators. Taking the transpose,

$$\ddot{B}_{i,k}^{\top} = \sum_{\mathbf{j}\in\mathbb{N}^N} \mathcal{F}^{-1} \circ \psi_{\mathbf{j}}^{c_{i,k}}(\mathbf{v}(\tilde{\xi}^{(k)})) \circ \mathcal{F} \circ (\rho_{i,k}a_k)(\tilde{x}, \tilde{\nu}_{i,k})\lambda_{\mathbf{j}}\psi_{\mathbf{j}}^{c_{i,k}}(\mathbf{u}(\tilde{x}^{(k)})) \tag{3.2.57}$$

Finally, we can truncate the infinite sum in (3.2.57) as before, and equations (3.2.54) and (3.2.57) then give us an formula for the transpose operator analogous to our earlier formula (3.2.46) for A.

Remark. As an alternative approach, we can develop this representation by writing $A^{\top}g$ locally as an oscillatory integral, similar to (3.2.1) but with the roles of the domain and codomain variables switched:

$$(A^{\top}g)(x) = \int e^{i(S(y,\xi)-x\cdot\xi)}a^{\top}(y,\xi)g(y)\,dy\,d\xi. \tag{3.2.58}$$

where $a^{\top}$ is an amplitude function. Separating out the first term in the Taylor expansion of the phase as before, we can write A^* locally as a sum of changes of coordinates and pseudodifferential operators, then perform a second-order Taylor approximation of the phase and expand the exponential using PSWFs. This leads once more to formulas (3.2.54) and (3.2.57).

Inverse

The major application of the transpose algorithm is computing the inverse of an elliptic Fourier integral operator. Since $(A^{-1})^{\top}$ is an FIO with canonical transformation χ, all

that is left is to compute its principal symbol. The rough intuition is that the principal symbol of A^{-1} should be the reciprocal of that of A, and the principal symbols of $(A^{-1})^\top$ and A^{-1} coincide. With our (discontinuous) trivializations of the half-density and Maslov bundles, this is exactly true. For $F \in I^m(\chi)$, denote by σ_F the principal symbol of F defined by (3.2.45).

Proposition 3.9. *If $A \in I^m(\chi)$ is elliptic, $\sigma_{A^{-1}} = \sigma_A^{-1}$.*

Lemma 3.10. *For any $A \in I^m(\chi)$, the principal symbols of A and $A^\top$ are identical.*

With these results, the inverse can be computed using the transpose algorithm applied to $(A^{-1})^\top$.

3.3 Examples

In this section, we compute a variety of FIOs using the algorithm described in section 3.2. All of these examples involve two-dimensional X and Y; while the algorithm applies equally well in higher dimensions, a few implementation details need to be worked out.

3.3.1 Radon transform (2D)

The 2D Radon transform is a prototypical FIO. In standard coordinates, every point is a caustic, so the use of coordinate changes is essential to avoid caustics.

Let θ be the angular coordinate on $\mathcal{S}^1$, and let $\hat{\theta} = (\cos\theta, \sin\theta) \in \mathbb{R}^2$. Then

$$Rf(s,\theta) = \int_{x\cdot\hat{\theta}=s} f(x)\, d\ell(x) = \frac{1}{2\pi}\int_{\mathbb{R}^2}\int e^{i\varsigma(x\cdot\hat{\theta}-s)} f(x)\, d\varsigma\, dx. \tag{3.3.1}$$

Letting $\varsigma, \vartheta, \xi$ be the dual variables to s, θ, x, the canonical relation of R is

$$\Lambda' = \left\{ (s,\theta,\varsigma,\vartheta;x,\xi) \;\middle|\; s = \pm x\cdot\frac{\xi}{|\xi|},\ \hat{\theta} = \pm\frac{\xi}{|\xi|},\ \varsigma = \pm|\xi|,\ \vartheta = \mp x\cdot\xi^\perp \right\} \subseteq (T^*Y \times T^*X)\setminus 0. \tag{3.3.2}$$

Λ' is a 1–2 map, equal to the disjoint union of two canonical graphs $\Lambda_\pm$, one for each choice of sign in (3.3.2). Hence we can write R as a sum of two graph FIOs $R_\pm$ associated with

their respective canonical graphs, with principal symbols equal to the restriction of σ_R to $\Lambda_\pm$.

From the canonical relation (3.3.2), we see that $\operatorname{rank} \partial(s,\theta)/\partial x = 1$, so that every point on $\Lambda_\pm$ is a caustic.

Principal Symbol Now let us return to the principal symbol of R. Writing R in the standard form (3.2.41) with $\theta = \sigma$, we have $n = 4$, $N = 1$, $\phi = \varsigma(x \cdot \hat{\theta} - s)$, so

$$Rf(s,\theta) = \frac{1}{2\pi} \int_{\mathbb{R}^2} \int_{\mathbb{R}} e^{i\varsigma(x\cdot\hat{\theta}-s)} f(x)\, d\varsigma\, dx = \frac{1}{(2\pi)^{3/2}} \int_{\mathbb{R}^2} \int_{\mathbb{R}} e^{i\varsigma(x\cdot\hat{\theta}-s)} \sqrt{2\pi} f(x)\, d\varsigma\, dx \tag{3.3.3}$$

The principal symbol is by (3.2.44)

$$\sigma = \sqrt{2\pi} \left(\frac{d_C}{|dx||d\xi|} \right)^{1/2} \exp\left[\frac{i\pi}{4} \operatorname{sgn} d_\varsigma^2 \phi \right]. \tag{3.3.4}$$

Since $d_\varsigma^2 \phi = 0$, the Maslov factor is 1. All that remains to compute d_C, the quotient volume form defined by $d_\varsigma \phi$. Now $d_\varsigma \phi$ maps the vector field $\frac{\partial}{\partial s}$ on $X \times Y \times \mathbb{R} = \{(x, s, \theta, \varsigma)\}$ to a unit-volume basis in its image, so since $|ds||d\theta||dx||d\varsigma|$ is a volume form on C, we get $d_C = |d\theta||dx||d\varsigma|$. Now we want to push d_C forward to Λ', where θ, ς form a polar coordinate representation for ξ (up to sign). Hence $|d\xi| = |\varsigma| \cdot |d\theta||d\varsigma|$, giving us

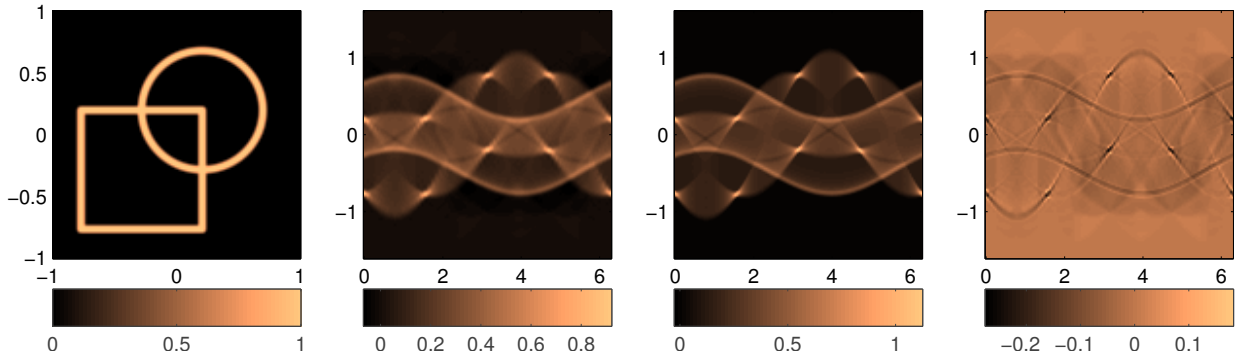
$$d_C = |d\theta|\, |dx|\, |d\varsigma| = \frac{|dx||d\xi|}{|\xi|}, \tag{3.3.5}$$

and finally

$$\sigma_R = \sqrt{2\pi} \left(\frac{d_C}{|dx||d\xi|} \right)^{1/2} = \sqrt{2\pi} \cdot |\xi|^{\,-1/2}. \tag{3.3.6}$$

As expected, σ_R is homogeneous of order $-1/2$, which is the order of R.

Example Output Figure 3.1 compares the Radon transform computed by the new algorithm with a naïve $O(N^3)$ reference implementation. Note that the asymptotic runtime of the new algorithm even with its greater complexity is $O(N^{2.5} \log N)$ (although $O(N^2 \log N)$ algorithms for computing the Radon transform are available).

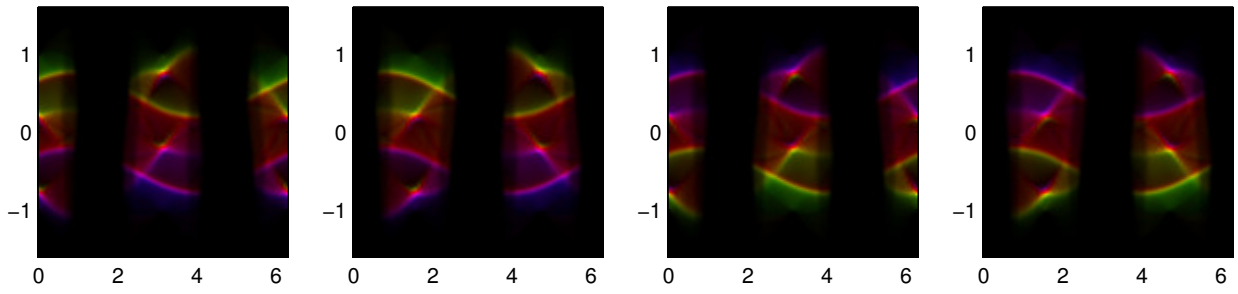


(a) f (input) (b) Computed $\mathcal{A}_\epsilon f$ (c) Reference Rf (d) Error, $\mathcal{A}_\epsilon f - Rf$

Figure 3.1: Comparison of generic FIO algorithm with reference Radon algorithm. θ and s are displayed on the x and y axes, respectively, as usual.

To handle caustics in $R = R_+ + R_-$, two changes of coordinates, $\beta_{1,(2)} \colon x_1 \mapsto x_1 + \frac{t}{2}x_2^2$ and $\beta_{2,(1)} \colon x_2 \mapsto x_2 + \frac{t}{2}x_1^2$ are used (see (3.2.38)). Since $\frac{dy}{dx}$ is identically zero, the identity coordinate change does not help here. Figure 3.2 shows the contributions to $R_+ f$ and $R_- f$ from each choice of coordinates.

Figures 3.3 and 3.4 shows the results of varying several parameters in the algorithm: the zero-padding factor, the number of frequency cones, and the amount of distortion in the coordinate changes (t in (3.2.38)). Increasing the zero-padding factor decreases periodicity artifacts that arise from the DFT. In the current setup, these artifacts appear as smooth errors outside the support of the true Rf. Increasing the number of frequency cones decreases the error in the approximation of the amplitude and phase (in the limit as the number of cones goes to infinity, the error $\|A - \mathcal{A}\|$ between the (ideal) computed operator $\mathcal{A}$ and the actual operator A goes to zero).



(a) R_+, $x_1 \mapsto x_1 + \frac{t}{2}x_2^2$ (b) R_+, $x_2 \mapsto x_2 + \frac{t}{2}x_1^2$ (c) R_-, $x_1 \mapsto x_1 + \frac{t}{2}x_2^2$ (d) R_-, $x_2 \mapsto x_2 + \frac{t}{2}x_1^2$

Figure 3.2: Contributions to Rf from each local integral representation: color represents the phase and brightness the magnitude of these complex-valued functions. Since f is real, R_+f and R_-f are complex conjugates.

3.3.2 Wave propagation

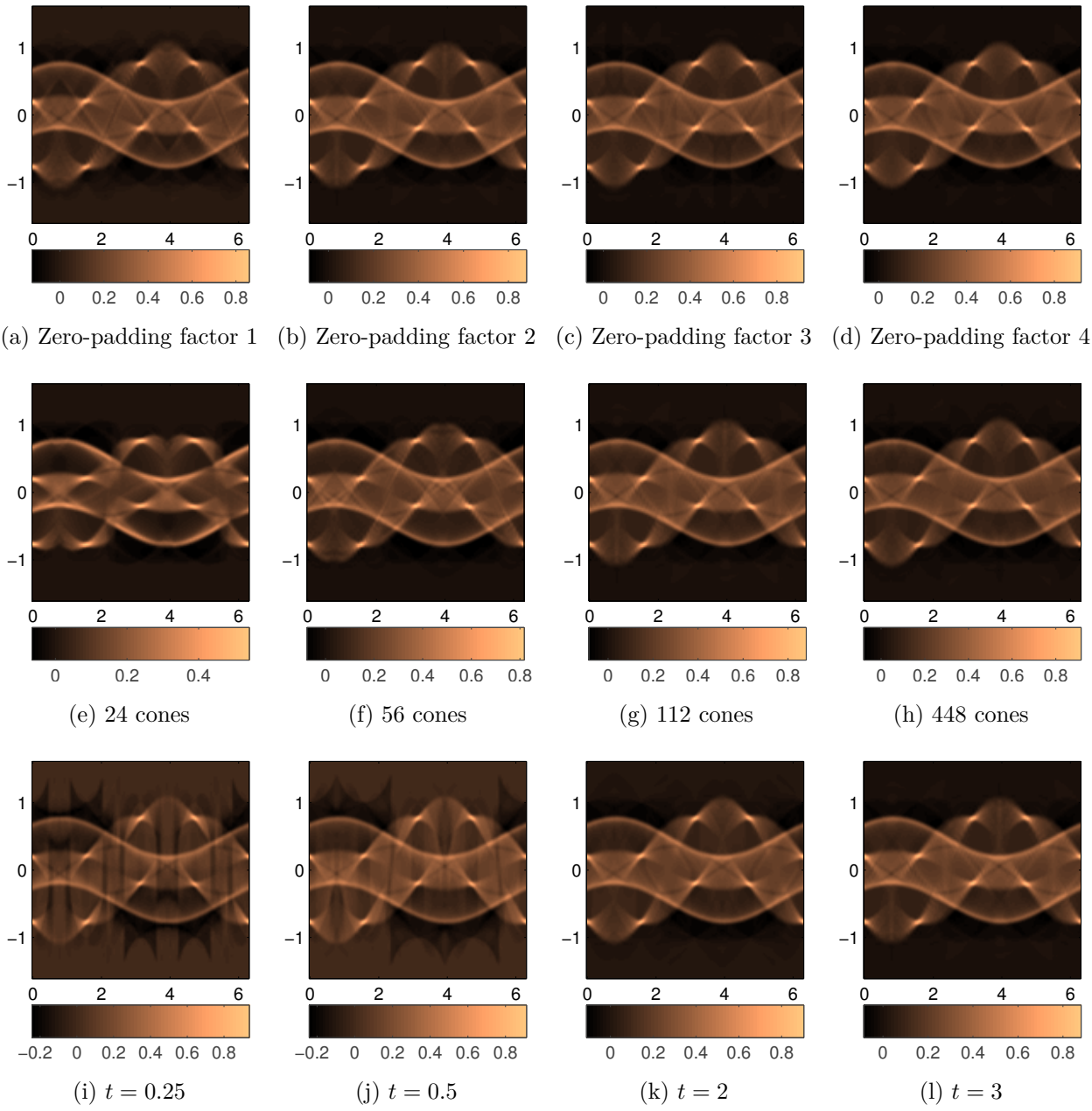
The initial value problem for the scalar wave equation with variable speed is a prime example of caustic formation:

$$\begin{cases} (\partial_t^2 - c^2\Delta)u = 0, & (x,t) \in X \times \mathbb{R}_+, \\ u(x,0) = g(x), \\ u_t(x,0) = h(x). \end{cases} \tag{3.3.7}$$

In this section, we will apply the algorithm to a propagator operator for a similar equation.

It is well-known that this solution operator is a sum of two FIOs (see for example [11]), and that the singularities of the initial data f follow the bicharacteristics of the wave operator $\Box_c = \partial_t^2 - c^2\Delta$. Furthermore, it is well-known that the bicharacteristics project to geodesics in X, and that a singularity at $(x,\xi) \in T^*X$ in the Cauchy data (g,h) propagates to two singularities in $u(\cdot,t)$ at $\Big(\gamma(x,\xi)(\pm t), \gamma'(x,\xi)(\pm t)\Big)$,where $\gamma(x,\xi)(s)$ is the unit-speed geodesic with $\gamma(0) = x$, $\gamma'(0) = \xi$.

Factoring $(\partial_t^2 - c^2\Delta)$ as $(\partial_t - iP)(\partial_t + iP)$ [17], where $P = \sqrt{-c^2\Delta}$ is a pseudodifferential square root, allows us to write any solution to (3.3.7) as a sum of solutions to the two

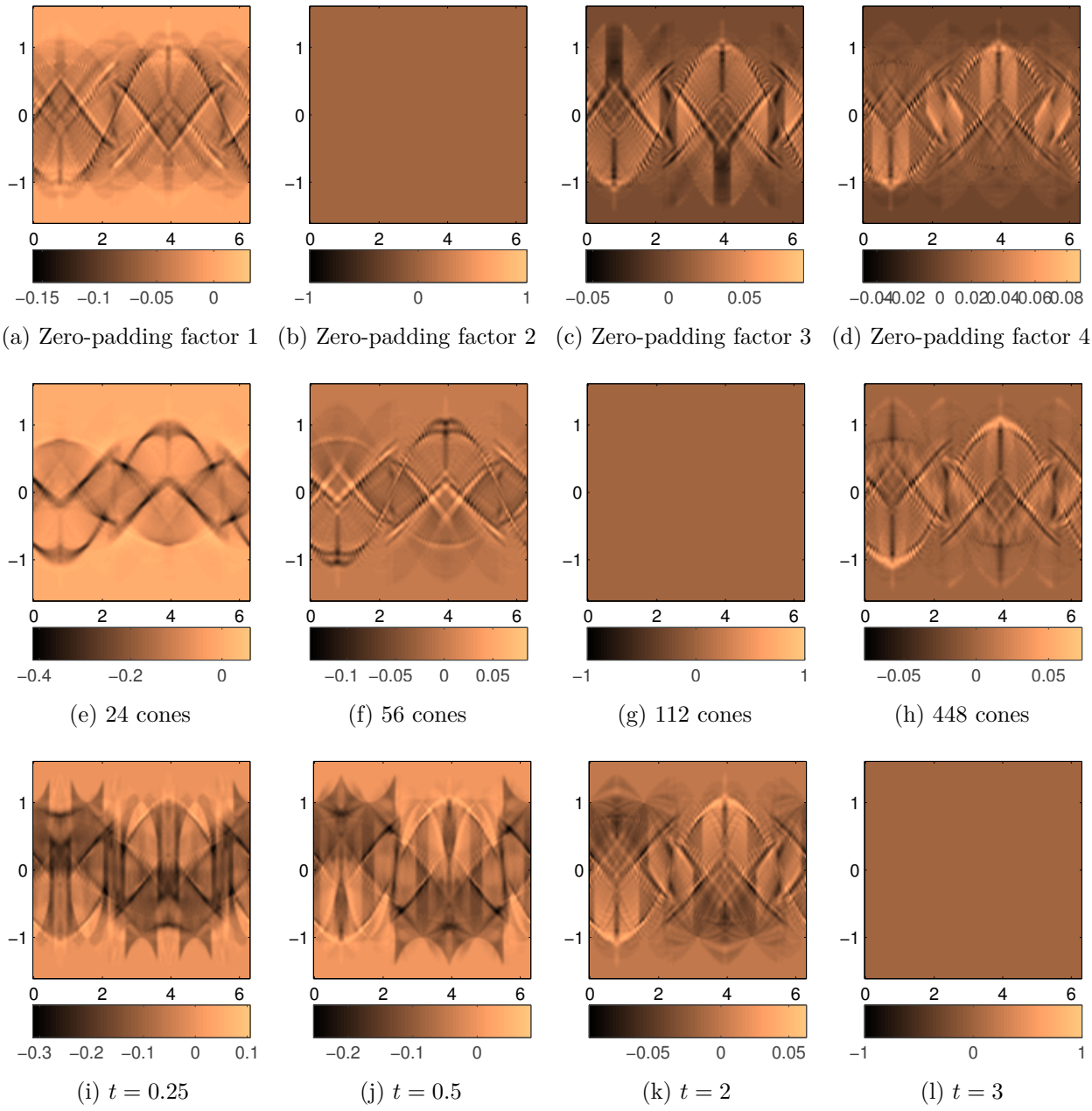


(a) Zero-padding factor 1 (b) Zero-padding factor 2 (c) Zero-padding factor 3 (d) Zero-padding factor 4

(e) 24 cones (f) 56 cones (g) 112 cones (h) 448 cones

(i) $t = 0.25$ (j) $t = 0.5$ (k) $t = 2$ (l) $t = 3$

Figure 3.3: Adjusting algorithm parameters. Top row: zero-padding factor for DFT; middle row: number of cones; bottom row: bending coefficient for diffeomorphisms. The base parameters were a zero-padding factor of 2, $t = 3$, and 112 cones.



(a) Zero-padding factor 1 (b) Zero-padding factor 2 (c) Zero-padding factor 3 (d) Zero-padding factor 4

(e) 24 cones (f) 56 cones (g) 112 cones (h) 448 cones

(i) $t = 0.25$ (j) $t = 0.5$ (k) $t = 2$ (l) $t = 3$

Figure 3.4: Relative errors in Figure 3.3, compared to $\mathcal{A}_\epsilon$ computed for the base parameters: a zero-padding factor of 2, 112 cones, and $t = 3$. Top row: zero-padding factor for DFT; middle row: number of cones; bottom row: bending coefficient for diffeomorphisms.

half-wave equations

$$\begin{cases} (\partial_t \pm iP)u = 0, & (x,t) \in X \times \mathbb{R}_+, \\ u(x,0) = f_\pm(x). \end{cases} \tag{3.3.8}$$

for suitable $f_\pm$ depending on g, h. The solution operator for (3.3.8) can be written as $S_\pm = e^{\mp itP}$. Restricting to T gives the operator $S_T = e^{iTP}$; this is what we will compute.

From the theory of propagation of singularities already mentioned, we know S_T has canonical relation

$$\chi(x,\xi) = \left(\gamma_{x,\xi}(T), \gamma'_{x,\xi}(T)\right). \tag{3.3.9}$$

For the principal symbol, since P has a real symbol, it is self-adjoint, so S_T is unitary: $S_T^* S_T = e^{-iTP^*} e^{iTP} = \mathrm{Id}$. By the principal symbol calculus, $\sigma_{S_T^* S_T} = |\sigma_{S_T}|^2$, so $|\sigma_{S_T}| = 1$. When T is small enough that there are no caustics, we can solve (3.3.8) using a Lax parametrix construction, taking S_+ of the form

$$(S_+ f)(t,x) = \int e^{i\phi(x,t,\xi)} a(x,t,\xi) \hat{f}(\xi)\, d\xi. \tag{3.3.10}$$

It is well known that ϕ satisfies an eikonal equation, and writing a in an asymptotic expansion $a \sim \sum a_{-j}$, where $a_{-j} \in S^{-j}$, and solving for S_+ gives a transport equation for the leading term a_0. This equation ensures that a_0 is real as long as ϕ is defined (i.e. before the formation of caustics). Evaluating at $t = T$ implies that the principal symbol of S_T without the Maslov factor, $\tilde{\sigma}_T = a d_C^{1/2} / (|dx|^{1/2} |d\xi|^{1/2})$, is real. Further, since $|\sigma_{S_T}| = 1$ and $\tilde{\sigma}_0 = 1$, we must have $\tilde{\sigma}_T = 1$ for all T.

The Maslov factor arising from the Lax parametrix construction is $-\operatorname{sgn}(\partial y_T/\partial x)^{-1}(\partial y_T/\partial \xi)$, where $y_T(x,\xi)$ is the y-component of the canonical transformation at (x,ξ). For T small enough (in particular before caustics have formed), $\operatorname{sgn}(\partial y_T/\partial x)^{-1}(\partial y_T/\partial \xi) = \operatorname{sgn}(\partial y_0/\partial x)^{-1}(\partial y_0/\partial \xi) = 1$ by continuity, giving a principal symbol of $\sigma_{S_T} = \exp[-i\pi/4]$. For larger T, we can find the principal symbol using Proposition 3.13; however, this has not yet been implemented.

Results To implement S_T, the canonical transformation χ was computed by solving the geodesic equation with RK4 time stepping. The geodesic computation is done in Christoffel

symbols allowing for anisotropic wave speeds.

Figures 3.6 to 3.10 compares the results of the algorithm with a standard reference algorithm that computes $S_{T,c}$ using a spectral method in space and RK4 time stepping. In the reference algorithm, P is approximated by its leading order term $c|D|$. The sound speed has a Gaussian low-speed lens (as in [16]) centered at the origin (Figure 3.5)

In Figure 3.6, solutions are calculated at $T = 0.3$, where no caustics have yet formed. The new algorithm compares quite favorably with the reference algorithm, without the high-frequency artifacts which arise most likely from applying P in the domain; the complex phases are correct.

It is worth noting that the FIO approach handles infinite domains naturally, without the need for absorbing boundary conditions, as Figure 3.7 shows; the pushforwards by the canonical transformation simply map escaping wavefronts outside the domain. Equally, periodic domains can be obtained by treating the codomain of the pushforwards as periodic, and performing no zero-padding, this is how Figure 3.6c is produced. Furthermore, Dirichlet or Neumann boundary conditions on arbitrary domains are likewise possible; the canonical transformation would then be modified so as to reflect wavefronts. This flexibility is one major advantage of the FIO approach.

In Figure 3.8, solutions are calculated at $T = 0.6$, showing the formation of caustics. Five coordinate choices were used here: $x_1 \mapsto x_1 \pm \frac{1}{2}x_2^2$ and $x_2 \mapsto x_2 \pm \frac{1}{2}x_1^2$, together with the identity. Figure 3.9 shows the contributions to $S_T f$ arising from each coordinate choice. These contributions are computed automatically. An algorithm for finding the correct Maslov factor is not yet implemented, so a phase difference is visible in the caustic in the error plot (d). However, the wavefronts are reconstructed correctly. Figure 3.10 shows the results when the coordinate changes are not sufficient to cover the caustics; in this figure, just two coordinate changes were used: $x_1 \mapsto x_1 + \frac{1}{2}x_2^2$ and $x_2 \mapsto x_2 + \frac{1}{2}x_1^2$.

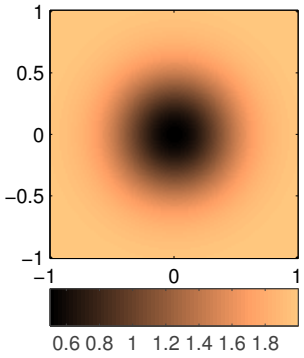
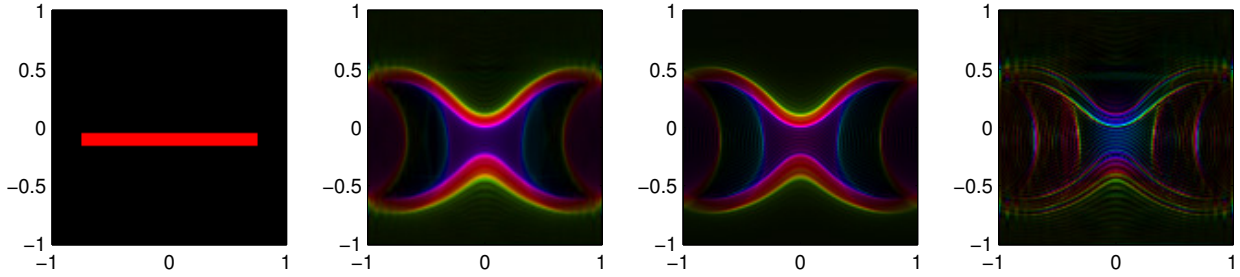
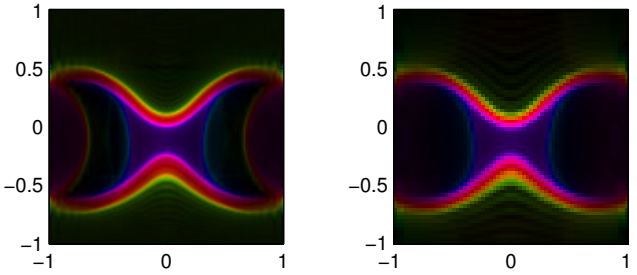


Figure 3.5: Lensing wave speed for examples.



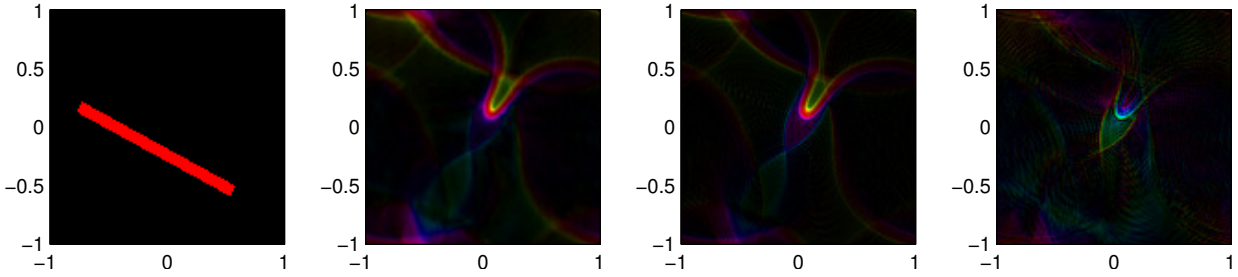
(a) f (input) (b) Computed $\mathcal{A}_\epsilon f$ (c) Reference $S_T f$ (d) Error, $\mathcal{A}_\epsilon f - Rf$

Figure 3.6: Comparison of generic FIO algorithm with reference half-wave propagator algorithm, no caustics. Color represents the phase and brightness the magnitude of these complex-valued functions.



(a) $\mathcal{A}_\epsilon f$, periodic grid (b) $\mathcal{A}_\epsilon f$, nonperiodic grid

Figure 3.7: Generic FIO algorithm with periodic and nonperiodic grid; nonperiodic grid conditions are handled automatically.



(a) f (input) (b) Computed $\mathcal{A}_\epsilon f$ (c) Reference $S_T f$ (d) Error, $\mathcal{A}_\epsilon f - Rf$

Figure 3.8: Comparison of generic FIO algorithm with reference half-wave propagator algorithm, with caustics present.

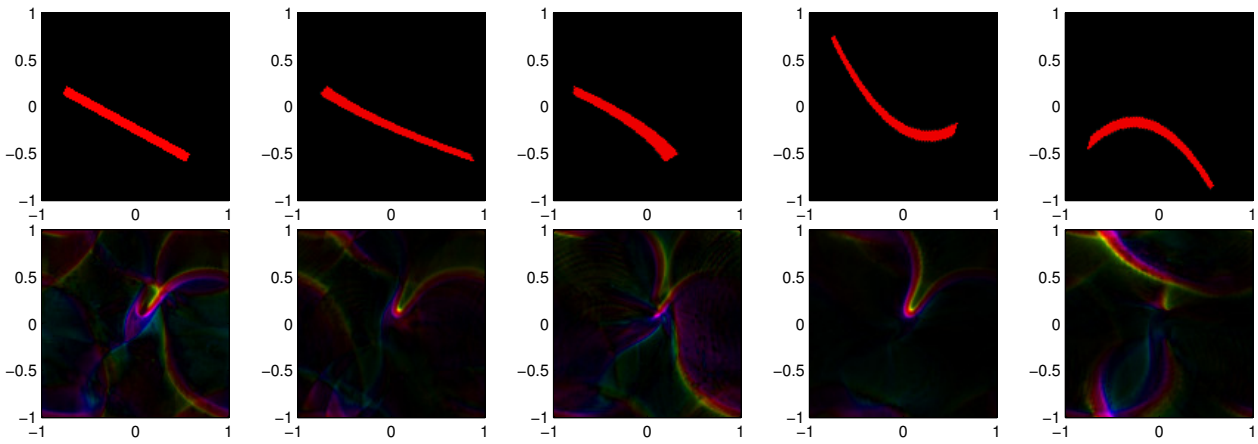
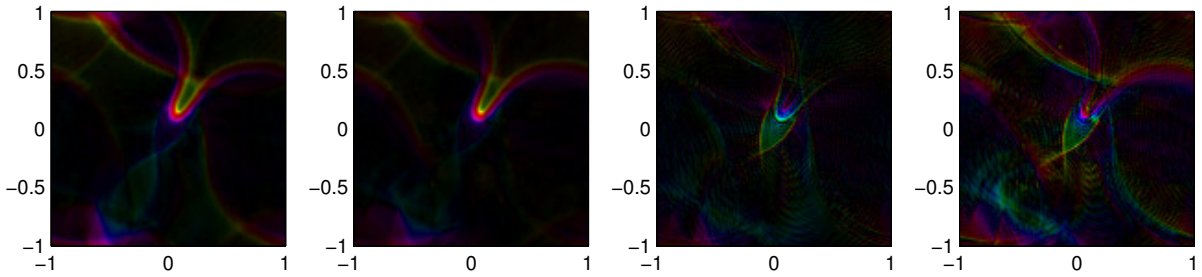


Figure 3.9: Coordinate changes and contributions of each to the computed $\mathcal{A}_\epsilon f$ for the setup of Figure 3.8.



(a) $\mathcal{A}_\epsilon f$, $Q = 5$ (b) $\mathcal{A}_\epsilon f$, $Q = 2$ (c) Error, $Q = 5$ (d) Error, $Q = 2$

Figure 3.10: Insufficient coordinate changes. (a) is computed with our algorithm with five coordinate choices. (b) is computed with just two coordinate choices; parts of the wavefront set are missing. (c) and (d) show the errors of each compared to the reference implementation.

3.3.3 SAR Reflection Operator

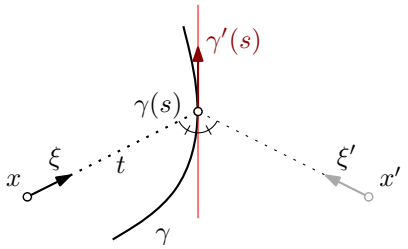
In synthetic aperture radar (SAR) imaging, a plane or satellite flies along a path γ, emitting radiation and measuring the echoes returned from the ground. Under a set of simplifying assumptions (flat terrain and an undirected beam), the forward (imaging) operator can be modeled by the circular Radon transform R_γ:

$$R_\gamma f(r,s) = \int_{|x-\gamma(s)|=r} f(x)\, d\ell. \tag{3.3.11}$$

Stefanov and Uhlmann [33] showed that R_γ is microlocally noninjective. After localizing γ near a certain s_0, and restricting to a sufficiently small set $X \subseteq \mathbb{R}^2 \setminus \gamma$, there is an FIO U such that $R_\gamma \circ U = R_\gamma$ microlocally, modulo smoothing operators. U can be thought of as reflection across the curve γ. The canonical relation of U is

$$\Lambda' = \left\{ (x',\xi',x,\xi) \;\middle|\; \exists s,t\colon\, x+t\xi = \gamma(s) = x'+t\xi',\ \dot\gamma(s)\cdot\xi = \dot\gamma(s)\cdot\xi',\ \dot\gamma(s)^\perp\cdot\xi = -\dot\gamma(s)^\perp\cdot\xi' \right\}. \tag{3.3.12}$$

Λ' is the graph of the map which sends singularities (x,ξ) to their mirror images (about the tangent line to γ) on the opposite side of γ:



In section 2.4, an algorithm for approximating the action of U was given, based on solving the wave equation forwards then backwards in time. In section 2.2, the principal symbol of U is shown to be 1, modulo a Maslov factor. Of course, for computational purposes, we require the precise Maslov factor:

Proposition 3.11. *The principal symbols of R_γ and U are:*

$$\sigma_{R_\gamma} = \sqrt{\frac{2\pi r}{|\xi||\sin\theta|}}, \tag{3.3.13}$$

$$\sigma_U = \exp\left[-\frac{i\pi}{4}\operatorname{sgn}\left(\xi \cdot \gamma''(s)\right)\right], \tag{3.3.14}$$

where θ is the angle between ξ and $\gamma'(s)$, as defined in section 2.2.3. Geometrically, the Maslov factor in σ_U indicates whether γ is concave toward or away from x at $\gamma(s)$.

The proof is given in section 3.5.

We now demonstrate the results of computing U with the current algorithm, compared to the reference wave-based algorithm described earlier in section 2.4. To compute the canonical transformation of U, we must locate the intersection points for the ray through (x, ξ) with γ. To do so, γ is represented as the zeros of a defining function $\Gamma(x)$, and then intersection points are found using an efficient recursive subdivision algorithm.

Figure 3.11 shows a simple example where γ is a parabola. The support of f is in the caustic-free region of Λ_U, so no coordinate changes are required. The new algorithm compares favorably to the reference wave-based algorithm visually, with much less high-frequency artifacts and a clear wavefront set.

In Figure 3.12, we compare $R_\gamma f$ to the $R_\gamma U f$ calculated by the new and reference algorithms. By definition of U, $R_\gamma f = R_\gamma U f$ modulo C^∞) for the exact operators. While the new algorithm has more smooth error above the support of the true $R_\gamma f$ as compared to the reference algorithm, the wavefront set is better preserved and there is less high-frequency noise. A better wavefront set can, however, be obtained with the reference algorithm by using a finer grid (Figure 3.12e); the current algorithm is independent of the codomain grid.

3.4 Implementation Details

In this section, we expand on some details of a freely-available implementation of the algorithm[1], which was implemented in object-oriented MATLAB with some C++ modules.

[1]Code is available by emailing the author at **pcaday@math.washington.edu**.

3.4.1 PSWF Computation

As described in section 3.2.1, our current implementation uses generalized PSWFs to expand the second-order approximated exponential functions $\exp[icu(x) \cdot v(\xi)]$, where $u, v \in \mathcal{B}^D$. As described in Slepian's work [26], a basis for $L^2(\mathcal{B}^D)$ can be formed from generalized PSWFs, which are products of radial parts $\phi_{N,n}^{(c)}(|z|)$ and spherical harmonics $Y_{D,N,n}(z/|z|)$. The radial parts are also eigenfunctions of a second-order Sturm-Liouville operator $L_{N,c}$.

Bases of spheroidal harmonics are computed following [2, (2.164)]. The radial functions $\phi_{N,n}^{(c)}$ are computed using Shkolinsky's algorithm [30], a variant of the algorithm described in [26]. Both methods are based on expanding the $\phi_{N,n}^{(c)}$ in a certain basis of polynomials $\{T_{N,m}^{(c)}\}_{m=0}^{\infty}$. Using the fact that they are also eigenfunctions of a second-order Sturm-Liouville operator L, it is shown that L is tridiagonal with respect to this basis. Truncating the expansion to the first K basis polynomials truncates L to a $K \times K$ matrix, giving K approximate eigenfunctions $\phi_{N,1}^{(c,K)}, \ldots, \phi_{N,K}^{(c,K)}$. Furthermore, the eigenvalues of the $\phi_{N,n}^{(c)}$ determine the eigenvalues of the generalized PSWFs with respect to the truncated Fourier transform [26].

For small K, increasing K increases the accuracy of the computed $\phi_{N,n}^{(c,K)}$ but for larger K the method becomes unstable experimentally. Stability is improved (and the domain of $\phi_{N,n}^{(c)}$ extended to $[0, \infty)$) by using the Bessel function expansion described in [26]. The error in $\phi_{N,n}^{(c,K)}$ can be estimated by measuring $\|\phi_{N,n}^{(c,2K)} - \phi_{N,n}^{(c,K)}\|$, and K may be successively doubled until this error estimate is satisfactory, subject to maintaining stability. As c increases, so does the number of significant eigenvalues of the truncated Fourier transform, requiring larger K.

3.4.2 Pushforwards

A key part of the practical implementation of the algorithm is the use of pushforwards to perform computations in global (x, ξ) coordinates on Λ rather than the local (y, ξ) coordinates. Suppose we are given a function f represented by its values at (regularly-spaced) grid points $(x_{\mathbf{i}}) = (x_{i_1,\ldots,i_n})_{1 \le i_k \le M_k}$, and a local diffeomorphism $T \colon X \to \mathbb{R}^n$. For each grid point

$x_{\mathbf{i}}$, $\mathbf{i} = (i_1, \ldots, i_n)$, let $\mathcal{N}_{\mathbf{i}}$ be the neighboring half-grid points $x_{i_1 \pm 1/2, \ldots, i_n \pm 1/2}$. Let $\mathcal{Q}_{\mathbf{i}}$ be the box spanned by $\mathcal{N}_{\mathbf{i}}$ and $\mathbf{1}_{\mathcal{Q}_{\mathbf{i}}}$ its indicator function. Now we push forward the half-grid points $\bigcup_{\mathbf{i}} \mathcal{N}_{\mathbf{i}}$ by T, and choose basis elements $U_{\mathbf{i}} \in L^\infty(Y)$ such that $R_{\mathbf{i}} \approx T_*(|DT|\mathbf{1}_{\mathcal{Q}_{\mathbf{i}}})$.

Since T is locally a diffeomorphism, if $Q_{\mathbf{i}}$ is sufficiently small (i.e., the grid is sufficiently fine), $T_*(|DT|\mathbf{1}_{\mathcal{Q}_{\mathbf{i}}}) = (T^{-1})^* \mathbf{1}_{\mathcal{Q}_{\mathbf{i}}} = \mathbf{1}_{(T^{-1})^* \mathcal{Q}_{\mathbf{i}}}$. Hence we can attempt to approximate this indicator function $\mathbf{1}_{(T^{-1})^* \mathcal{Q}_{\mathbf{i}}}$ by indicator functions of simpler polyhedra.

For $n = 2$, if the quadrilateral described by the corners $T(\mathcal{N}_{\mathbf{i}})$ is not self-intersecting, we may take $\mathcal{R}_{\mathbf{i}}$ to be the indicator function of this quadrilateral. In the self-intersecting case, the quadrilateral describes a union of two triangles, and we may take $R_{\mathbf{i}}$ to be the sum of their indicator functions. However, since $d\chi$ is everywhere nonsingular, each image quadrilateral is non self-intersecting for sufficiently fine grids.

For $n > 2$ dimensions, the situation is more complex, as vertices of $\mathcal{N}_{\mathbf{i}}$ lying on a common face are not in general coplanar after applying T. One solution is to break each box $\mathcal{N}_{\mathbf{i}}$ into simplices and let $R_{\mathbf{i}}$ be the indicator function of the union of the corresponding simplices spanned by the images $T(\mathcal{N}_{\mathbf{i}})$. In order to ensure that there are no overlaps or gaps between adjacent $R_{\mathbf{i}}$, adjoining faces of neighboring boxes must have compatible decompositions into simplices. Another, more symmetric, solution is to fix a single simplex decomposition D (represented as an collection of simplices on an abstract set of points) and its images under the symmetry group $\mathcal{K}_n$ of the n-cube. Then, take $R_{\mathbf{i}}$ to be the average of the indicator functions corresponding to each image decomposition. In general, either choice of $R_{\mathbf{i}}$ is only a reasonable approximation to $T_*(|DT|\mathbf{1}_{\mathcal{Q}_{\mathbf{i}}})$ if the distorted hypercube with corners $T(\mathcal{N}_{\mathbf{i}})$ is non-self-intersecting.

After choosing the $R_{\mathbf{i}}$, we apply a filter ϕ_S and sample at the codomain grid points:

$$T_* f(y_{\mathbf{i}}) \approx \sum_{\mathbf{i}} (R_{\mathbf{i}} * \phi_S)(y_{\mathbf{i}}). \tag{3.4.1}$$

In the current 2D implementation, an elliptical filter ϕ_S was used: ϕ_S is the indicator of the ellipse $E = \{(y_1/\Delta y_1)^2 + (y_2/\Delta y_2)^2 \leq 1\}$, where Δy_i is the grid spacing in each coordinate. With this filter, $(R_{\mathbf{i}} * \phi_S)(y_{\mathbf{i}})$ is the area of the intersection of the polygon(s) $\operatorname{supp} R_{\mathbf{i}}$ with

the ellipse centered at $y_{\mathbf{i}}$.

The runtime of this algorithm is $O(\ell + m')$, where ℓ is the number of grid points in the domain, and m' the sum of the grid points covered by each $R_{\mathbf{i}} * \phi_S$. If $\kappa = \sup_{y \in Y} |T^{-1}(y)|$ is the maximum number of preimages of points in Y under T, then $m' = O(\ell + \kappa m)$, where m is the number of codomain grid points: the $\mathbf{i}^{\text{th}}$ box in the domain grid maps to $O(1 + v_{\mathbf{i}}/\Delta y)$ points, where $v_{\mathbf{i}}$ is the volume of the image of the box (assuming the box is non-self-intersecting), inside the codomain grid, and Δy the volume of a box in the codomain grid. The $O(1)$ factor comes from the support of the filter ϕ_S. Summing over all domain boxes, the total volume of box images is at most $\kappa \operatorname{vol}(Y)$, and $\kappa \operatorname{vol}(Y)/\Delta y = \kappa m$. The total runtime is on the order of the number of points in the codomain (counted with multiplicity), $O(\ell + \kappa m)$.

3.4.3 Low Frequencies

As has already been mentioned, the leading-term approximations make the current algorithm effectively a high-frequency algorithm. In particular, near the origin, the low angular resolution of the Cartesian DFT becomes problematic, especially for FIOs of nonpositive degree. If only the high frequencies of the output are desired, a high-pass cutoff can be applied to the input f prior to the algorithm; but, as the cutoff is not coordinate-invariant, it must be applied after pulling back to $\tilde{x}^{(i)}$ coordinates. This is a rather unsatisfactory solution, however.

Another option is the use a nonuniform discrete Fourier transform to sample lower frequencies more densely. While there are good general-purpose nonuniform DFT algorithms available, we employ the *pseudopolar FFT* of [3], which computes $\hat{f}(\xi)$ on a *pseudopolar* grid. The pseudopolar grid is similar to a regular polar grid except with concentric squares in place of circles, and constant spacing of slopes instead of angles. The runtime for the pseudopolar FFT is $O(N^2 \log N)$ time with a small constant, and $O(N^2 \log N)$ time to invert using an iterative least-squares method (with the constant depending on the desired accuracy ϵ). As well, Levi [22] alludes to an $O(N^2 \log N)$ direct inversion method.

In our current implementation, the pseudopolar FFT is available as an option, but is relatively slow, and an optimized implementation of the pseudopolar FFT is needed.

A third possibility is to apply a high-pass filter $\rho(\xi)$ as before, then approximate the low-frequency component $(1-\rho(D))f$ of the input as a sum of plane waves whose frequencies lie on a pseudopolar grid, and apply the FIO to each directly. This approach is successful with some simple FIOs, but not in general.

3.5 Proofs

3.5.1 Sufficiency of Quadratic Coordinate Changes

Proposition 3.7. *For all $\gamma_0 \in \Lambda'$, there exist $J \subsetneq \{1,\dots,n\}$ and $i \in \{1,\dots,n\} \setminus J$ such that $\det \partial y/\partial \tilde{x}^{(i,J)}(\gamma_0) \neq 0$, where $\tilde{x}^{(i,J)} = \beta_{i,J}(x)$.*

For the proof of proposition 3.7, we use the following lemma:

Lemma 3.12. *Let $E = T_{\gamma_0}(T^*X \times T^*Y)$, and let $L = T_{\gamma_0}\Lambda'$. There is a Lagrangian subspace $M \subseteq E$ with $L \cap M = \{0\}$ of the form*

$$M_J = \operatorname{span}\left\{ \left\{ \frac{\partial}{\partial x^i} \;\middle|\; i \notin J \right\} \cup \left\{ \frac{\partial}{\partial \xi_j} \;\middle|\; j \in J \right\} \cup \left\{ \frac{\partial}{\partial \eta_1}, \dots, \frac{\partial}{\partial \eta_n} \right\} \right\}, \quad J \subsetneq \{1,\dots,n\}. \tag{3.5.1}$$

Proof. Let σ be the symplectic form on E, and let $M = \operatorname{span}\{\partial/\partial\eta_1, \dots, \partial/\partial\eta_n\}$. Since Λ' is the graph of χ, the projection $\pi\colon \Lambda' \to T^*X$ is a diffeomorphism, which implies $L \cap M = \{0\}$. Now let us enlarge M in two ways: $M_1 = \operatorname{span}\{M, \frac{\partial}{\partial x^1}\}$ and $M_2 = \operatorname{span}\{M, \frac{\partial}{\partial \xi_1}\}$. Both are isotropic. Suppose $L \cap M_1, L \cap M_2$ are both nontrivial, and pick nonzero $v_1 = m_1 + a\frac{\partial}{\partial x^1}$, $v_2 = m_2 + b\frac{\partial}{\partial \xi_1}$ in $L \cap M_1$, $L \cap M_2$, respectively. Then

$$\sigma(v_1, v_2) = \sigma(m_1, m_2) + a\sigma(\tfrac{\partial}{\partial x^1}, m_2) + b\sigma(m_1, \tfrac{\partial}{\partial \xi_1}) + ab\sigma(\tfrac{\partial}{\partial x^1}, \tfrac{\partial}{\partial \xi_1}) = ab. \tag{3.5.2}$$

Since L is Lagrangian, $\sigma(v_1, v_2) = 0$, so one of a, b is zero, and therefore one of v_1, v_2 lies in $L \cap M$, a contradiction. Hence at least one of $L \cap M_1$, $L \cap M_2$ is $\{0\}$. Replace M by one

of the M_i for which is true. Repeating the process with $\frac{\partial}{\partial x^k}$, $\frac{\partial}{\partial \xi_k}$ for $k = 2, \dots, n$, we obtain the M described in (3.5.1). There is one point to verify: that $J \neq \{1, \dots, n\}$. Because Λ' is conic, the radial vector $\sum(\frac{\partial}{\partial \xi_k}(\xi_0)_k + \frac{\partial}{\partial \eta_k}(\eta_0)_k)$ lies in $L = T_{\gamma_0}\Lambda'$, where $\gamma_0 = (x_0, \xi_0, y_0, \eta_0)$. Hence M cannot contain all the $\frac{\partial}{\partial \xi_k}$, and $J \neq \{1, \dots, n\}$. □

Proof of Proposition 3.7. By lemma 3.12, there is some M of the form (3.5.1) transversal to L. Of all these, choose one where J is minimal; i.e. no $M_{J'}$ with $J' \subsetneq J$ is transversal to L. Let $I = \{1, \dots, n\} \setminus J$. Now M_J is the kernel of $d\pi_J$, where π_J is the projection of $T^*X \times T^*Y$ onto (x_J, ξ_I, y). $L \cap M_J = \emptyset$ is equivalent to $d\pi_J|_{\Lambda'}$ being bijective; so (x_J, ξ_I, y) are local coordinates for Λ'. Since (x, ξ) are also local coordinates, this is equivalent to the non-singularity of $[\frac{\partial y}{\partial x_I} \frac{\partial y}{\partial \xi_J}]$ at γ_0.

Now we turn to $\frac{\partial y}{\partial \tilde{x}}$ for the $\beta_{i,J}$ map (where i is to be determined). Applying (3.2.37) to the $\beta_{i,J}$,

$$\frac{\partial y}{\partial \tilde{x}_k} = \begin{cases} \frac{\partial y}{\partial x_k}, & k \in J \\ \frac{\partial y}{\partial x_k} + \xi_i \frac{\partial y}{\partial \xi_k}, & k \notin J. \end{cases} \tag{3.5.3}$$

Choose some $i \in I$ such that $(\xi_0)_i \neq 0$. Such an i exists because otherwise the radial vector $\rho = \sum(\frac{\partial}{\partial \xi_k}(\xi_0)_k + \frac{\partial}{\partial \eta_k}(\eta_0)_k)$ of $T_{\gamma_0}\Lambda'$ would lie in M_J. Since J is minimal, for $j \in J$, $\frac{\partial y}{\partial x_j}$ is a linear combination of $\{\frac{\partial y}{\partial x_k} : k \in I\}$. Hence, because $[\frac{\partial y}{\partial x_I} \frac{\partial y}{\partial \xi_J}]$ is nonsingular at γ_0, so is $\frac{\partial y}{\partial \tilde{x}}$. □

3.5.2 Principal Symbol Calculus

It is well-known that the principal symbol of the composition of graph FIOs is given intuitively by multiplying them (up to factors coming from the Maslov and half-density bundle). The following proposition states how composition works with the trivialization of these bundles we are using.

Proposition 3.13. *Let $A \in I^{m_1}(\chi_1)$, $B \in I^{m_2}(\chi_2)$ be graph FIOs associated with canonical transformations $\chi_1 \colon T^*X \to T^*Y$, $\chi_2 \colon T^*Y \to T^*Z$. Then the principal symbol of the*

composition BA is

$$\sigma_{BA} = \sigma_A \sigma_B \exp\left[\frac{i\pi}{4}\left(\operatorname{sgn}(Q_A + Q_B) - \operatorname{sgn} Q_A - \operatorname{sgn} Q_B\right)\right], \tag{3.5.4}$$

where $Q_A = -\partial y(x,\eta)/\partial\eta$ and $Q_B = \partial y(z,\eta)/\partial\eta$ for any choice of local coordinates y on Y such that (x,η), resp. (z,η) are local coordinates on the graph of χ_1, resp. χ_2.

Proof. Our starting point is the standard proof of FIO composition, Theorem 1.11. Working microlocally near some $\gamma \in T^*X$, we can write the composition BA as an FIO plus a smoothing operator (which we disregard as usual), after performing a change of coordinates $y \mapsto \mathfrak{y}/H$, $H = |(\theta_1, \theta_2)|$, where the θ_i are the frequency variables for the phases of A and B.

To simplify matters, we have restricted our attention to graph FIOs. Choose local coordinates x, y, z on X, Y, Z respectively, and write $(x(y,\eta), \xi(y,\eta))$ and $(z(y,\eta), \zeta(y,\eta))$ for the coordinate functions of χ_1^{-1} (defined locally near $\chi_1(x_0,\xi_0)$) and χ_2 respectively. Assuming that $\det \partial x/\partial y(\chi_1(\gamma)) \neq 0$, we have that (x,η) are local coordinates on Λ_1 near γ. Similarly, if $\det \partial z/\partial y(\chi_1(\gamma)) \neq 0$ then (z,η) are local coordinates on Λ_2 near $\chi_1(\gamma)$. Given the choice of x, z coordinates, we can always choose y generically so that both determinants are nonzero.

Let $S_1(x,\eta) = \eta \cdot y^{(1)}(x,\eta)$, $S_2(z,\eta) = \eta \cdot y^{(2)}(z,\eta)$. Then $\phi_1(x,y,\eta) = y\cdot\eta - S_1(x,\eta)$ is a phase function parameterizing Λ_1, and likewise $\phi_2(y,z,\eta) = S_2(z,\eta) - y\cdot\eta$ parameterizes Λ_2. Hence there exist amplitudes $a \in S^{m_1}$, $b \in S^{m_2}$ such that A and B can be written modulo smoothing operators as

$$(Af)(y) = (2\pi)^{-n} \iint e^{i[y\cdot\eta - S_1(x,\eta)]} a(y,\eta) f(x)\, dx\, d\eta \tag{3.5.5}$$

$$(Bg)(z) = (2\pi)^{-n} \iint e^{i[S_2(z,\eta) - y\cdot\eta]} b(y,\eta) g(y)\, dy\, d\eta \tag{3.5.6}$$

for $\mathrm{WF}(f)$, $\mathrm{WF}(g)$ supported in small neighborhoods of γ_0 and $\chi_1(\gamma_0)$, respectively. Formally composing A and B, we proceed as in the proof of Theorem 1.11, substituting $y = \mathfrak{y}/H$ and multiplying the amplitude by a cutoff $\rho(\eta_1,\eta_2)$ near $\{\eta_1 = 0\} \cup \{\eta_2 = 0\}$, yielding

$$BAf(z) = (2\pi)^{-2n} \iiiint e^{i[S_2(z,\eta_2) - S_1(x,\eta_1) + \mathfrak{y}\cdot(\eta_1-\eta_2)/H]} a\left(\frac{\mathfrak{y}}{H}, \eta_1\right) b\left(\frac{\mathfrak{y}}{H}, \eta_2\right) H^{-n} f(x)\, dx\, d\eta_1\, d\mathfrak{y}\, d\eta_2. \tag{3.5.7}$$

The critical set of the phase is

$$C_\phi = \{(x, z, \mathfrak{y}, \eta_1, \eta_2) \mid \eta_1 = \eta_2,\ \mathfrak{y}/H = \partial_{\eta_1} S_1 = \partial_{\eta_2} S_2\}\,. \tag{3.5.8}$$

On C_ϕ, we write $\eta = \eta_1 = \eta_2$. To compute the half-density and Maslov factor components of the principal symbol, we need the second derivative matrix $d_{(\mathfrak{y},\theta_1,\theta_2)} d\phi$ on C_ϕ. Using Lemma 3.14, we find

$$d_{(\mathfrak{y},\eta_1,\eta_2)} d_{(x,z,\mathfrak{y},\eta_1,\eta_2)}\phi = \begin{bmatrix} 0 & 0 & 0 & \frac{I}{H} & -\frac{I}{H} \\ -\partial_x\partial_{\eta_1} S_1 & 0 & \frac{I}{H} & Q_A - \frac{1}{H^2}(y\eta^\top + \eta y^\top) & 0 \\ 0 & \partial_z\partial_{\eta_2} S_2 & -\frac{I}{H} & 0 & Q_B + \frac{1}{H^2}(y\eta^\top + \eta y^\top) \end{bmatrix} \tag{3.5.9}$$

where $Q_A = -\partial y(z,\eta)/\partial\eta$, $Q_B = \partial y(x,\eta)/\partial\eta$.

Half-density factor: To find the half-density factor, we need the quotient volume d_C on C_ϕ. Since $d_{(\mathfrak{y},\eta_1,\eta_2)}$ maps $(\frac{\partial}{\partial x}, \frac{\partial}{\partial z}, \frac{\partial}{\partial \eta_2})$ to a basis in $\mathbb{R}^{3n}$ with volume $H^{-n}|\partial_x\partial_{\eta_1} S_1||\partial_x\partial_{\eta_2} S_2|$,

$$\begin{aligned} d_C &= |dY|\,|d\eta_1|\,|H|^{\,n}\,|\partial_x\partial_{\eta_1} S_1|\,|\partial_x\partial_{\eta_2} S_2| && (3.5.10)\\ &= |dy|\,|d\eta_1|\,|H|^{\,2n}\,|\partial_x\partial_{\eta_1} S_1|\,|\partial_x\partial_{\eta_2} S_2| && (3.5.11)\\ &= |dx|\,|d\xi|\,|H|^{\,2n}\,|\partial_x\partial_{\eta_1} S_1|\,|\partial_x\partial_{\eta_2} S_2|\,. && (3.5.12) \end{aligned}$$

Similarly, we can find from (3.5.6) the half-density terms in the principal symbols of A and B. Denoting them d_{C_1}, d_{C_2} respectively, we have $d_{C_1} = |dx||d\xi||\partial_x\partial_{\eta_1} S_1|$, $d_{C_2} = |dy||d\eta||\partial_x\partial_{\eta_2} S_2|$.

Maslov factor: Denoting similarity by $\sim$, starting from Lemma 3.14

$$d^2_{(\mathfrak{y},\eta_1,\eta_2)}\phi \sim \begin{bmatrix} 0 & I & -I \\ I & Q_A & 0 \\ -I & 0 & Q_B \end{bmatrix} \sim \begin{bmatrix} 0 & I & -I \\ I & Q_A - \epsilon I & 0 \\ -I & 0 & Q_B + \epsilon I \end{bmatrix} \tag{3.5.13}$$

for any ϵ. Choosing ϵ so that $Q_A - \epsilon I$ is nonsingular, we apply further similarity transforms to achieve

$$d^2_{(\mathfrak{y},\eta_1,\eta_2)}\phi \sim \begin{bmatrix} -Q_A + \epsilon I & 0 & 0 \\ 0 & Q_A - \epsilon I & 0 \\ 0 & 0 & Q_A + Q_B \end{bmatrix} \tag{3.5.14}$$

Hence $\operatorname{sgn} d^2_{(\mathfrak{y},\eta_1,\eta_2)}\phi = \operatorname{sgn}(Q_A + Q_B)$; by comparison, the signatures in the Maslov factors for A and B are just Q_A and Q_B. Combining this with (3.5.12) and the amplitude of (3.5.7), we have the following principal symbol for BA:

$$\sigma_{BA} = |H|^{-n} ab\rho \sqrt{|H|^{2n} \left|\partial_x \partial_{\eta_1} S_1\right| \left|\partial_x \partial_{\eta_2} S_2\right|} \exp\left[\frac{\pi i}{4} \operatorname{sgn}(Q_A + Q_B)\right]. \tag{3.5.15}$$

No powers of 2π occur in (3.5.15) as the power of 2π occurring in (3.5.7) is exactly the right one for $3n$ frequency variables in n dimensions. We can also drop ρ from (3.5.15) since this modifies σ_{BA} only by a term in $S^{-\infty}$. Since

$$\sigma_A = a\sqrt{\left|\partial_x \partial_{\eta_1} S_1\right|} \exp\left[-\frac{\pi i}{4} \operatorname{sgn} Q_A\right], \qquad \sigma_B = b\sqrt{\left|\partial_x \partial_{\eta_1} S_1\right|} \exp\left[\frac{\pi i}{4} \operatorname{sgn} Q_B\right], \tag{3.5.16}$$

this establishes the formula for the principal symbol microlocally near γ. Since γ was arbitrary, we are done. □

In the proof, we needed a computational lemma that we will use again later on for computing the Maslov factor for the SAR reflection operator:

Lemma 3.14. *Let $\phi_1(x, y, \theta_1)$, $\phi_2(y, z, \theta_2)$ be phase functions, and $\phi(x, \mathfrak{y}, z, \theta_1, \theta_2) = \phi_1(x, \mathfrak{y}/H, \theta_1) + \phi_2(\mathfrak{y}/H, z, \theta_2)$, where $H = |(\theta_1, \theta_2)|$. Let $d^2_{y,\theta_i}\phi_i = \left[\begin{smallmatrix} A_i & B_i^\top \\ B_i & C_i \end{smallmatrix}\right]$, let $A = H^{-2}(A_1 + A_2)$, and let $M_i = H^{-1}\theta_i y^\top$. Then*

$$d^2_{(\mathfrak{y},\theta_1,\theta_2)}\phi = \begin{bmatrix} A & * & * \\ \frac{B_1}{H} - M_1 A & C_1 - \frac{B_1 M_1^\top + M_1 B_1^\top}{H} + M_1 A M_1^\top & * \\ \frac{B_2}{H} - M_2 A & -\frac{B_2 M_1^\top + M_1 B_2^\top}{H} + M_2 A M_1^\top & C_2 - \frac{B_2 M_2^\top + M_2 B_2^\top}{H} + M_2 A M_2^\top \end{bmatrix}, \tag{3.5.17}$$

where the $*$ entries are given by symmetry. Furthermore, letting $\sim$ denote similarity,

$$d^2_{(\mathfrak{y},\theta_1,\theta_2)}\phi \sim \begin{bmatrix} A_1 + A_2 & B_1^\top & B_2^\top \\ B_1 & C_1 & 0 \\ B_2 & 0 & C_2 \end{bmatrix}. \tag{3.5.18}$$

Proof. The Hessian (3.5.17) is by direct computation. Applying the similarity transform $Vd^2_{(\mathfrak{y},\theta_1,\theta_2)}V^\top$ gives (3.5.18), where

$$V = \begin{bmatrix} H & & \\ M_1 & I & \\ M_2 & & I \end{bmatrix}. \tag{3.5.19}$$

□

Note that Proposition 3.13 holds for any FIOs A and B whose canonical relations are locally canonical graphs. An easy corollary (and the application here) is the principal symbol of the inverse:

Proposition 3.9. *If $A \in I^m(\chi)$ is elliptic, $\sigma_{A^{-1}} = \sigma_A^{-1}$.*

Proof. Let $\gamma \in T^*X$. Since χ is a symplectomorphism, it is a local diffeomorphism and therefore χ^{-1} is locally a graph. Now, apply Proposition 3.13 to the composition $\mathrm{Id} = A^{-1}A$. The principal symbol of the identity is 1 (equal to its principal symbol as a pseudodifferential operator), so

$$1 = \sigma_A \sigma_{A^{-1}} \exp \frac{\pi i}{4} \left[\operatorname{sgn}(Q_{A^{-1}} + Q_A) - \operatorname{sgn}(Q_{A^{-1}}) - \operatorname{sgn}(Q_A) \right]. \tag{3.5.20}$$

Since A^{-1} is associated with the canonical relation χ^{-1}, we have $Q_{A^{-1}} = -Q_A$ ($y(x,\eta) = y(z,\eta)$ in the terminology of the proof of Proposition 3.13). Hence $\sigma_A \sigma_{A^{-1}} = 1$. □

Lemma 3.10. *For any $A \in I^m(\chi)$, the principal symbols of A and $A^\top$ are identical.*

Proof. This is immediate since the kernels differ only in the ordering of x and y, which does not affect the value of the principal symbol. □

3.5.3 SAR Reflection Operator Principal Symbol

Proposition 3.11. *The principal symbols of R_γ and U are:*

$$\sigma_{R_\gamma} = \sqrt{\frac{2\pi r}{|\xi||\sin\theta|}}, \tag{3.3.13}$$

$$\sigma_U = \exp\left[-\frac{i\pi}{4} \operatorname{sgn}\left(\xi \cdot \gamma''(s)\right) \right], \tag{3.3.14}$$

where θ is the angle between ξ and $\gamma'(s)$, as defined in section 2.2.3. Geometrically, the Maslov factor in σ_U indicates whether γ is concave toward or away from x at $\gamma(s)$.

Proof of formula for σ_{R_γ}. For $\theta \in \mathbb{R}$, define $\hat{\theta} = (\cos\theta, \sin\theta)$. First, define

$$Rf(r,p) = \int_0^{2\pi} r f(p + r\hat{\theta})\, d\theta. \tag{3.5.21}$$

Applying the inverse Fourier transform in r and its inverse, $\mathcal{F}_r\mathcal{F}_r^{-1}$, inside the integral,

$$Rf(r,p) = \frac{1}{2\pi}\int_0^{2\pi}\int_{\mathbb{R}}\int_{\mathbb{R}} e^{i\rho(r'-r)} r' f(p + r'\hat{\theta})\, d\rho\, dr'\, d\theta. \tag{3.5.22}$$

Changing from polar coordinates,

$$Rf(r,p) = \frac{1}{2\pi}\int_{\mathbb{R}}\int_{\mathbb{R}^2} e^{i\rho(|x-p|-r)} f(x)\, d\rho\, dx. \tag{3.5.23}$$

Restricting to γ gives us

$$R_\gamma f(r,s) = (2\pi)^{-3/2}\int_{\mathbb{R}}\int_{\mathbb{R}^2} e^{i\rho(|x-\gamma(s)|-r)}\sqrt{2\pi} f(x)\, d\rho\, dx. \tag{3.5.24}$$

ϕ is a phase function (in particular, smooth) for x restricted to $\mathbb{R}^2 \setminus \gamma$, and

$$dd_\rho\phi = \begin{bmatrix} 1 & \frac{x-\gamma(s)}{|x-\gamma(s)|}\cdot\gamma'(s) & -\frac{x-\gamma(s)}{|x-\gamma(s)|} & 0 \end{bmatrix}. \tag{3.5.25}$$

Hence ϕ is also nondegenerate. So $d\rho$ maps $\frac{\partial}{\partial r}$ to a unit volume basis, implying $d_C = |d\rho||ds||dx|$. Letting $\xi = d_x\phi$, we can compute

$$\left|\frac{\partial\xi}{\partial(\rho,s)}\right| = \left|\det\begin{bmatrix}\frac{\xi}{|\xi|} & \frac{\rho\gamma'(s)}{|x-\gamma(s)|}\end{bmatrix}\right| = \left|\frac{\rho}{r}\frac{\xi}{|\xi|}\cdot\gamma'(s)\right| = \frac{|\xi|}{r}\,|\sin\theta.| \tag{3.5.26}$$

Hence $d_C = |\rho r^{-1}\sin\theta||dx||d\xi|$, and the Maslov factor is 1 since $d_\rho^2\phi = 0$. Therefore, the principal symbol of R_γ is $\sqrt{2\pi r/|\xi||\sin\theta|}$, as claimed. As expected, it has degree $-1/2$.

Proof of formula for σ_U. Let Σ_L', Σ_R' be open sets in $T^*(\mathbb{R}^2\setminus\gamma)$ (chosen as the W_j in 2.2.1) such that rays through elements $(x,\xi) \in \Sigma_L'$, resp. Σ_R' intersect γ exactly once from the left, resp. right, and $\Sigma_R' = \mathcal{C}(\Sigma_L')$. Choose a subset $\Sigma_L \subset \overline{\Sigma_L} \subset \Sigma_L'$ and let $\Sigma_R = \mathcal{C}(\Sigma_L)$. Let α_L, α_R be bump functions on $T^*\mathbb{R}^2$ supported in Σ_L', Σ_R' respectively, and equal to 1 on Σ_L and Σ_R respectively. Define $R_{\gamma,L} = R_\gamma\alpha_L(x,D)$ and similarly for $R_{\gamma,R}$.

By Proposition 3.13, the principal symbol of $R_{\gamma,R}$ is equal to σ_{R_γ} on Σ_R, which is bounded away from zero. Therefore $R_{\gamma,R}$ is elliptic on Σ_R and hence has a microlocal inverse $R_{\gamma,R}^{-1}$ such that $R_{\gamma,R}^{-1}R_{\gamma,R} = \mathrm{Id}$ microlocally on Σ_R. By Proposition 3.9, the symbol of $R_{\gamma,R}^{-1}$ is the reciprocal of that of $R_{\gamma,R}$. For f with $\mathrm{WF}(f) \subseteq \Sigma_L$, we have $R_\gamma f = R_{\gamma,L}f = R_{\gamma,R}R_{\gamma,R}^{-1}R_{\gamma,L}f$ modulo C^∞, hence by the definition of the reflection operator (2.2.3), $U = R_{\gamma,R}^{-1}R_{\gamma,L}$ microlocally.

With this in hand, we can compute the principal symbol of U via Proposition 3.13. Since the principal symbols of $R_{\gamma,L}$ and $R_{\gamma,R}^{-1}$ are reciprocals, all that remains is the Maslov factor. Instead of computing it from Proposition 3.13, we express $R_{\gamma,L}$, $R_{\gamma,R}^{-1}$ in a different form, using 3.5.24 to write $R_{\gamma,\square}$ ($\square$ indicating L or R) to leading order as

$$R_{\gamma,\square}f(r,s) = (2\pi)^{-3/2}\int_{\mathbb{R}}\int_{\mathbb{R}^2} e^{i\rho(|x-\gamma(s)|-r)}\sqrt{2\pi}\tilde{\alpha}_\square(x,r,s,\rho)f(x)\,d\rho\,dx \\ \mod I^{-3/2}(\mathbb{R}_+\times I\times\mathbb{R}^2). \quad (3.5.27)$$

where $\tilde{\alpha}_\square$ are some order zero amplitudes. (Specifically, choose $\tilde{\alpha}_\square$ so that the restriction of $\tilde{\alpha}$ to the critical set C_ϕ is equal to $\alpha_\square$ after both are pulled back to Λ'; then the principal symbols of both sides of (3.5.27) are identical.) Now, composing $R_{\gamma,R}^{-1}\circ R_{\gamma,L}$ and proceeding as in the proof of Proposition 3.13, we obtain a Maslov factor $\exp[\frac{i\pi}{4}\operatorname{sgn}Q]$ where $Q = d^2_{\mathfrak{y},\rho_L,\rho_R}\phi$ (the subscripts on ρ and other variables indicate the FIOs $R_{\gamma,R}^{-1}$ or $R_{\gamma,L}$ they originate from). The signature of Q can be obtained from Lemma 3.14; if $\varphi_\square(r,s,\rho_\square,x)$ is the phase in (3.5.27), then we can compute

$$d^2_{r,s,\rho}\varphi_\square = \begin{bmatrix} 0 & 0 & 1 \\ 0 & (\dagger_\square) & -\xi_\square\cdot\gamma'(s) \\ 1 & -\xi_\square\cdot\gamma'(s) & 0 \end{bmatrix}, \qquad (\dagger_\square) = -\xi_\square\cdot\gamma''(s) - \rho_\square\frac{(\frac{\xi_\square}{|\xi_\square|}\cdot\gamma'^{\perp}(s))^2}{|\xi|}. \quad (3.5.28)$$

Applying Lemma 3.14, we have

$$d^2_{\mathfrak{y},\rho_L,\rho_R} \sim \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & (\dagger_L)+(\dagger_R) & -\frac{\sigma_L}{\rho_L} & -\frac{\sigma_R}{\rho_R} \\ 1 & -\frac{\sigma_L}{\rho_L} & 0 & 0 \\ 1 & -\frac{\sigma_R}{\rho_R} & 0 & 0 \end{bmatrix} \quad (3.5.29)$$

Applying similarity transformations we can reduce the matrix to

$$d^2_{\mathfrak{y},\rho_L,\rho_R} \sim \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & (\dagger_L) + (\dagger_R) & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \tag{3.5.30}$$

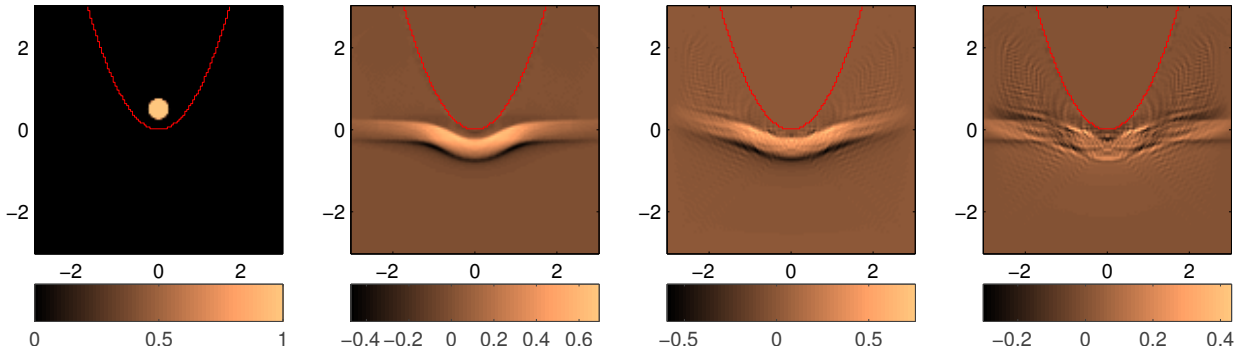
which has signature equal to $\operatorname{sgn}[(\dagger_L) + (\dagger_R)]$. On the critical set of the phase of composition $R^{-1}_{\gamma_R} R_{\gamma_L}$, $\rho_L = -\rho_R$. Also, $\gamma''(s)$ is perpendicular to $\gamma'(s)$, and $\frac{\xi_L}{|\xi_L|} \cdot \gamma'^{\perp}(s) = -\frac{\xi_R}{|\xi_R|} \cdot \gamma'^{\perp}(s)$. Hence the second terms in $(\dagger_L)$, $(\dagger_R)$ cancel, while the first terms are identical, leaving $(\dagger_L + \dagger_R) = -2\xi_L \cdot \gamma''(s)$. Hence we have a Maslov factor of $\exp[-\frac{\pi i}{4} \operatorname{sgn}(\xi_L \cdot \gamma''(s))]$. Dropping the L gives us the formula for σ_U. □

3.6 Future work

In this section we cover a few of the many avenues for future work on this problem. On the implementation side, the current program could be extended to handle 3D Fourier integral operators. We would like to apply the program to further graph FIOs, particularly generalized Radon transforms, including the geodesic X-ray transform. As well, the implementation of the pseudopolar FFT algorithm is currently a major bottleneck, and should be replaced by a C implementation, possibly incorporating the direct inversion algorithm alluded to by Levi [22].

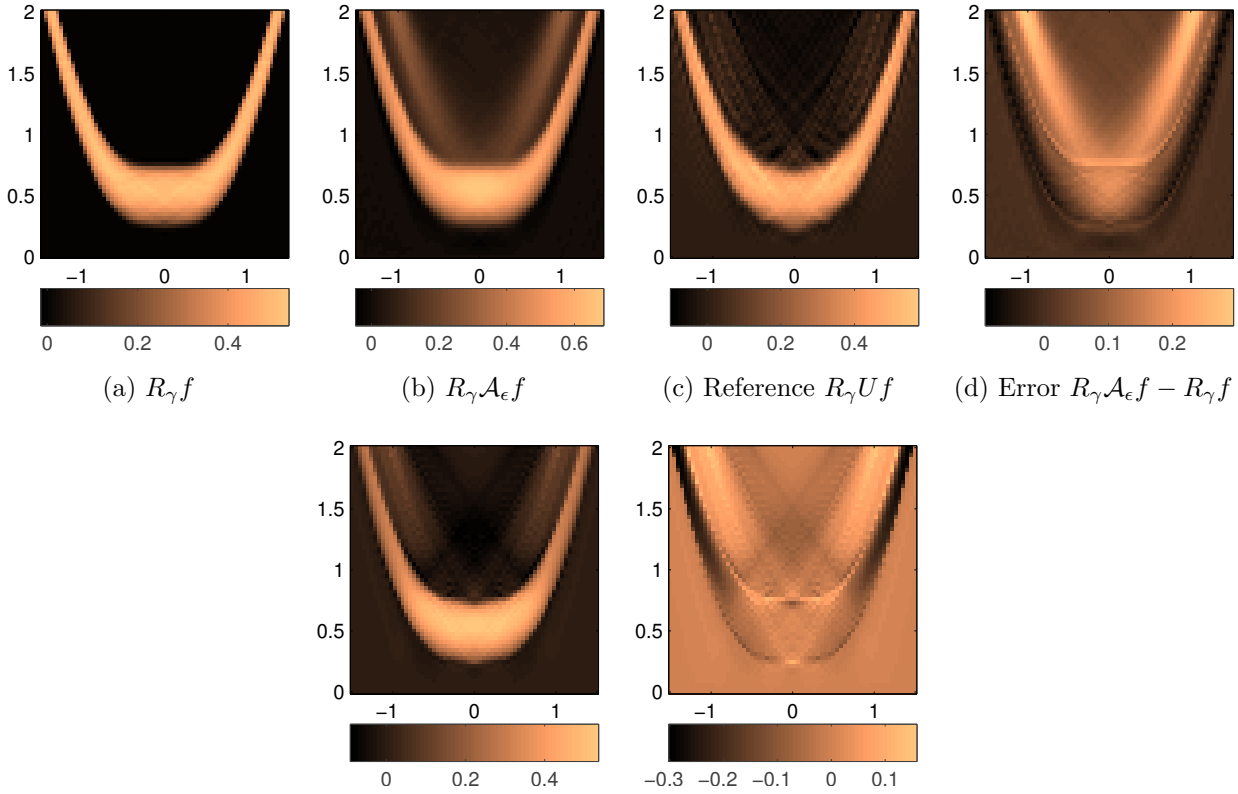
On the algorithmic side, the precomputed data requires a substantial amount of memory, $O(L^{(3n-1)/2})$. It would be desirable to reduce the storage requirements. One possibility would be to leverage the known sparsity results for FIOs with respect to wave packet frames, by computing the wave packet coefficients of the FIO applied to a single wave packet (or group of wave packets) and compressing them, calculating only the most significant coefficients. With the coefficient matrix, applying an FIO in the local form (3.2.1) reduces to a wave packet transform, a sparse matrix-vector multiplication, and then the inverse wave packet transform.

As well, we would like to calculate the original operator A rather than the approximation $\mathcal{A}_\epsilon$, which is approximate only to leading order. A_ϵ is an FIO associated with the same canonical transformation as A, so if A is elliptic, then $A = A(\mathrm{Id} - A^{-1}\mathcal{A}_\epsilon)$, and $\mathrm{Id} - A^{-1}\mathcal{A}_\epsilon$ is a pseudodifferential operator, which could potentially be inverted by a Neumann series or other methods.



(a) f (input) (b) $\mathcal{A}_\epsilon f$ (c) Reference Uf (d) Error, $\mathcal{A}_\epsilon f - Uf$

Figure 3.11: Comparison of the new algorithm with a reference algorithm for computing the reflection operator U.



(a) $R_\gamma f$ (b) $R_\gamma \mathcal{A}_\epsilon f$ (c) Reference $R_\gamma U f$ (d) Error $R_\gamma \mathcal{A}_\epsilon f - R_\gamma f$

(e) Reference $R_\gamma U f$, finer grid (f) Error $R_\gamma U f - R_\gamma f$

Figure 3.12: Applying the operator R_γ to each function in Figure 3.11, for points on γ with $|x| \leq 1.5$. For the exact reflection operator U, we would have $R_\gamma f = R_\gamma U f \bmod C^\infty$. In (e), Uf is computed using the reference algorithm on a grid half the size and spacing of that used for (c).

BIBLIOGRAPHY

[1] F. Andersson, M.V. de Hoop, and H. Wendt, Multiscale discrete approximation of Fourier integral operators, Multiscale Model. Simul. 10 (2012), no. 1, 111–145.

[2] K. Atkinson and W. Han, "Spherical Harmonics and Approximations on the Unit Sphere: An Introduction." Lecture Notes in Mathematics 2044. Berlin-Heidelberg: Springer-Verlag 2012. DOI 10.1007/978-3-642-25983-8_2

[3] A. Averbuch, R. Coifman, D. Donoho, M. Elad, M. Israeli, Fast and accurate polar Fourier transform. Appl. Comput. Harmon. Anal. 21 (2006), no. 2, 145–167.

[4] G. Bao and W. W. Symes, Computation of pseudo-differential operators, SIAM J. Sci. Comput. 17 (1996), no. 2, 416–429.

[5] A. Bonami and A. Karoui, Uniform bounds of prolate spheroidal wave functions and eigenvalues decay, C. R. Acad. Sci. Paris, Ser. I 352 (2014) 229–234.

[6] A. Bonami and A. Karoui, Spectral decay of the Sinc kernel operators and approximation by Prolate Spheroidal Wave Functions, submitted for publication (2014), arXiv:1012.3881v3.

[7] E. Candès and L. Demanet, Curvelets and Fourier integral operators, C. R. Math. Acad. Sci. Paris 336 (2003), no. 5, 395–398.

[8] E. Candès and L. Demanet, Curvelets, Warpings, and Optimal Representations of Fourier Integral Operators, technical report, 2002. Available online at http://math.mit.edu/icg/papers/CurveletsFIO_proof.pdf.

[9] E. Candès, L. Demanet, D. Donoho, and L. Ying, Fast Discrete Curvelet Transforms, Multiscale Model. Simul. 5 (3), 861–899 (2006).

[10] E. Candès, L. Demanet, and L. Ying, A Fast Butterfly Algorithm for the Computation of Fourier Integral Operators. Multiscale Model. Simul 7 (4), 1727–1750 (2009).

[11] E. Candès, L. Demanet, and L. Ying, Fast Computation of Fourier Integral Operators, SIAM J. Sci. Comput. 29 (6) 2464–2493 (2007).

[12] E. J. Candès and D. L. Donoho, Curvelets: A Surprisingly Effective Nonadaptive Representation for Objects With Edges, in Curves and Surface Fitting, Saint Malo, 1999; Albert Cohen, Christophe Rabut, and Larry L. Schumnaker, eds. Nashville: Vanderbilt University Press, 2000. ISBN 0-8265-1357-3.

[13] E. Cordero, K. Gröchenig, and F. Nicola, Approximation of Fourier integral operators by Gabor multipliers. J. Fourier Anal. Appl. 18 (2012), no. 4, 661–684.

[14] A. A. Duchkov, F. Andersson, and M. V. de Hoop, Discrete almost-symmetric wave packets and multiscale geometrical representation of (seismic) waves. IEEE Trans. Geoscience and Remote Sensing 48 (2010), no. 9, 3408–3423.

[15] M. V. de Hoop, H. Smith, G. Uhlmann, R. D. van der Hilst, Seismic imaging with the generalized Radon transform: a curvelet transform perspective. Inverse Problems 25 (2009), no. 2, 025005, 21 pp.

[16] M. V. de Hoop, G. Uhlmann, A. Vasy, and H. Wendt, Multiscale discrete approximations of Fourier integral operators associated with canonical transformations and caustics. Multiscale Model. Simul. 11, 566–585 (2013)

[17] L. Demanet and L. Ying, Discrete symbol calculus, SIAM Rev. 53 (2011), no. 1, 71–104.

[18] J.J. Duistermaat, Fourier Integral Operators, Modern Birkhäuser Classics, doi:10.1007/978-0-8176-8108-1_2

[19] K. Guo and D. Labate, Representation of Fourier integral operators using shearlets. J. Fourier Anal. Appl. 14 (2008), no. 3, 327–371.

[20] L. Hörmander, Analysis of Linear Partial Differential Operators, vols. III–IV. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 274–275. Springer-Verlag, Berlin, 1994.

[21] H. J. Landau and H. Widom, Eigenvalue distribution of time and frequency limiting, Journal of Mathematical Analysis and Applications 77 (1980), 469–481.

[22] O. Levi, Direct Exact Inverse Pseudo-Polar FFT and Radon Transform Using Orthogonalizing Weights, Summer School at Inzell (2014) (presentation slides).

[23] D. Slepian, H. O. Pollak, Prolate spheroidal wave functions, Fourier analysis and uncertainty. I. Bell System Tech. J. 40 (1961), 43–63.

[24] H. J. Landau, H. O. Pollak, Prolate spheroidal wave functions, Fourier analysis and uncertainty. II. Bell System Tech. J. 40 (1961), 65–84 .

[25] H. J. Landau, H. O. Pollak, Prolate spheroidal wave functions, Fourier analysis and uncertainty. III. The dimension of the space of essentially time- and band-limited signals. Bell System Tech. J. 41 (1962), 1295–1336.

[26] D. Slepian, Prolate spheroidal wave functions, Fourier analysis and uncertainty. IV. Extensions to many dimensions; generalized prolate spheroidal functions. Bell System Tech. J. 43, (1964) no. 6, 3009–3057.

[27] M. Rieffel, Quantization and C^*-Algebras, Contemporary Mathematics (167) 1994.

[28] Y. Safarov, A symbolic calculus for Fourier integral operators, in Geometric and Spectral Analysis, Contemporary Mathematics 630 (2014), 275–290.

[29] A. Seeger, C. Sogge, and E. Stein, Regularity properties of Fourier integral operators, Ann. of Math. (2) 134 (1991), no. 2, 231–251.

[30] Y. Shkolinsky, Prolate spheroidal wave functions on a disc—integration and approximation of two-dimensional bandlimited functions. Appl. Comput. Harmon. Anal. 22 (2007), no. 2, 235–256.

[31] H. Smith, A parametrix construction for wave equations with $C^{1,1}$ coefficients, Ann. Inst. Fourier (Grenoble) 48 (1998), no. 3, 797–835.

[32] H. Smith, A Hardy space for Fourier integral operators, J. Geom. Anal. 8 (1998), no. 4, 629–653.

[33] P. Stefanov and G. Uhlmann, Is a curved flight path in SAR better than a straight one? SIAM J. Appl. Math. (73) 1596–1612.

[34] H. Widom, Asymptotic behavior of the eigenvalues of certain integral equations. II. Arch. Rational Mech. Anal. 17 (1964) 215–229.