Chapter 8

Microlocal analysis of imaging

In this chapter we consider that the receiver and source positions (x_r, x_s) are restricted to a vector subspace $\Omega \subset \mathbb{R}^6$, such as $\mathbb{R}^2 \times \mathbb{R}^2$ in the simple case (!) when $x_{r,3} = x_{s,3} = 0$. We otherwise let x_r and x_s range over a continuum of values within Ω, which allows us to take Fourier transforms of distributions of x_r and x_s within Ω. We call Ω the *acquisition manifold*, though in the sequel it will be a vector subspace for simplicity[1].

8.1 Preservation of the wavefront set

In this section we show that, in simple situations, the operators F and F^* defined earlier have inverse kinematic behaviors, in the sense that the respective mappings of singularities that they induce are inverses of one another.

The microlocal analysis of F and F^* hinges on the mathematical notion of wavefront set that we now introduce. We start by describing the characterization of functions with (or without) singularities.

Nonsingular functions are infinitely differentiable, hence have Fourier transforms that decay "super-algebraically".

Lemma 1. *Let $u \in C_0^\infty(\mathbb{R}^n)$. Then, for all $N > 0$,*

$$|\widehat{u}(k)| \leq C_N (1 + |k|^2)^{-N}.$$

[1]The Fourier transform can still be defined for functions on general manifolds, but this involves patches, a partition of unity, and many distractions that we prefer to avoid here. The microlocal theory presented here carries over mostly unchanged to the manifold setting.

Proof. In the Fourier integral defining $\widehat{u}(k)$, insert N copies of the differential operator $L = \frac{I-\Delta_x}{1+|k|^2}$ by means of the identity $e^{ix\cdot k} = L^N e^{ix\cdot k}$. Integrate by parts, pull out the $(1+|k|^2)^{-N}$ factor, and conclude by boundedness of $u(x)$ and all its derivatives. □

A singularity occurs whenever the function is not locally C^∞.

Definition 2. *Let $u \in \mathcal{S}'(\mathbb{R}^n)$, the space of tempered distributions on $\mathbb{R}^n$. We define the singular support of u as*

$$\textit{sing supp}(u) = \{x \in \mathbb{R}^n : \textit{there does not exist an open}$$

$$\textit{neighborhood } V \textit{ of } x \textit{ such that } u \in C^\infty(V)\}.$$

In addition to recording where the function is singular, we also want to record the *direction* in which it is singular. This idea gives rise to the notion of wavefront set, that we now build up to.

"Direction of singularity" is associated to lack of decay of the Fourier transform in the same direction. For this purpose it is useful to consider sets invariant under rescaling, i.e., cones.

Definition 3. *A set X is said to be a conic neighborhood of a set $Y \subset \mathbb{R}^n$ if*

- *X is open;*
- *$Y \subset X$;*
- *$\xi \in X$ implies $\lambda\xi \in X$ for all $\lambda > 0$.*

Definition 4. *The singular cone of a tempered distribution $u \in \mathcal{S}'(\mathbb{R}^n)$ is the set*

$$\Gamma(u) = \{\eta \in \mathbb{R}^n\backslash\{0\} : \textit{there does not exist a conic}$$

$$\textit{neighborhood } W \textit{ of } \eta \textit{ such that for } \xi \in W,$$

$$|\widehat{u}(k)| \le C_N(1+|k|^2)^{-N}, \qquad \textit{for all } N > 0.\}$$

The intuition is that $\Gamma(u)$ records the set of directions ξ in which the Fourier transform of u decays slowly. The reason for formulating the definition in negative terms is that we want the resulting cone to be closed, i.e., isolated directions in which the Fourier transform would accidentally decay quickly, while decay is otherwise slow along an open set of nearby directions, do not count.

The construction of $\Gamma(u)$ is global – a direction is labeled singular without regard to the locations x where there are singularities that may have contributed to the direction being labeled singular. To go further we need to localize the construction.

Lemma 2. *Let $\phi \in C_0^\infty(\mathbb{R}^n)$ and $u \in \mathcal{S}'(\mathbb{R}^n)$. Then $\Gamma(\phi u) \subset \Gamma(u)$.*

Definition 5. *Let $x \in X \subset \mathbb{R}^n$. We call singular fiber at x the set*

$$\Gamma_x(u) = \{\bigcap_\phi \Gamma(\phi u) : \phi \in C_0^\infty(X), \phi(x) \neq 0\}.$$

The idea of this definition is that we localize u by means of multiplication with a smooth function ϕ of arbitrarily small support, then consider the smallest resulting singular cone. The definition can equivalently be formulated by means of a family of smooth indicators whose support converges toward the singleton $\{x\}$ in the sense of sets:

$$\Gamma_x(u) = \{\lim_{j\to\infty} \Gamma(\phi_j u) : \phi_j \in C_0^\infty(X), \phi_j(x) \neq 0, \operatorname{supp} \phi_j \to \{x\}\}.$$

Note that $\Gamma_x(u)$ is empty if u is smooth (C^∞) at x.

The wavefront set then consists of the union of the singular cones, taken as fibers over the singular support.

Definition 6. *Let $u \in \mathcal{S}'(\mathbb{R}^n)$. The wavefront set of u is the set*

$$WF(u) = \{(x, \xi) \in \mathbb{R}^n \times (\mathbb{R}^n \backslash \{0\}) : \xi \in \Gamma_x(u)\}.$$

The set $\mathbb{R}^n \times (\mathbb{R}^n \backslash \{0\})$ is also abbreviated as $T^*\mathbb{R}^n \backslash 0$.

Because this definition localizes the singularites of u both in position and in direction, it is the basis for *microlocal* analysis.

In this context, the product $\mathbb{R}^n \times \mathbb{R}^n$ is the cotangent bundle of $\mathbb{R}^n$, and denoted $T^*\mathbb{R}^n$. The reason for viewing each singular cones (each fiber) as a subset of the cotangent space at x rather than the tangent space at x will be apparent in the sequel as we study the behavior of $WF(u)$ under change of variables.

Notice that the projection of $WF(u)$ on the first $\mathbb{R}^n$ (relative to x) is the singular support sing supp(u), while the projection of $WF(u)$ on the second $\mathbb{R}^n$ (relative to ξ) is the singular cone $\Gamma(u)$.

In the context of imaging, we will use the same notation as introduced above and let ξ be the wave-vector variable, dual to position $x \in \mathbb{R}^n$ in model space. Usually, $n = 3$. Data space is indexed by the base variables $(x_s, x_r, t) \in \Omega \times \mathbb{R} \equiv \mathbb{R}^{n_d}$, usually with $n_d = 5$. We denote by (ξ_r, ξ_s, ω) the corresponding dual variables, i.e., the independent variables of the Fourier transform of a function of (x_s, x_r, t). The wavefront set of a distribution $d(x_r, x_s, t)$ in $\mathbb{R}^{n_d}$ is then

$$WF(d) = \{(x_r, x_s, t; \xi_r, \xi_s, \omega) \in T^*\mathbb{R}^{n_d}\backslash 0 : (\xi_r, \xi_s, \omega) \in \Gamma_{(x_r,x_s,t)}(d)\}.$$

We will also consider wavefront sets of functions in the product space $\mathbb{R}^n \times \mathbb{R}^{n_d} = \mathbb{R}^{n+n_d}$ of model space and data space, such as the distributional integral kernel K of F. We have

$$WF(K) = \{(x_r, x_s, t, x; \xi_r, \xi_s, \omega, \xi) \in T^*\mathbb{R}^{n+n_d}\backslash 0 : (\xi_r, \xi_s, \omega, \xi) \in \Gamma_{(x_r,x_s,t,x)}(K)\}. \tag{8.1}$$

In a product space, a wavefront set has the additional interpretation of being a *relation*. The role of $WF(K)$, where K is the kernel of F, is that it provides a way of predicting $WF(Fm_1)$ from $WF(m_1)$ using a simple composition rule.

In order to see $WF(K)$ as a relation, however, two minor modifications need to be made to the definition (8.1). First, the variables first need to be re-arranged as $((x, \xi), (x_r, x_s, t; \xi_r, \xi_s, \omega))$, in order to be seen as elements of the product space $T^*\mathbb{R}^n \times T^*\mathbb{R}^{n_d}$. Second, the sign of the first co-variable ξ needs to be flipped, so that we instead consider the "wavefront prime" $WF'(K)$. In short,

$$((x, -\xi), (x_r, x_s, t; \xi_r, \xi_s, \omega)) \in WF'(k) \quad \Leftrightarrow \quad (x_r, x_s, t, x; \xi_r, \xi_s, \omega, \xi) \in WF(K)$$

We call $WF'(K)$ the wavefront relation of the operator whose kernel is K.

More generally, the operations that define relations are as follows.

- The composition of a relation $C \subset X \times Y$ and a set $S \subset X$ is the set

$$C \circ S = \{y \in Y : \text{ there exists } x \in S \text{ such that } (x, y) \in C\}.$$

- The composition of two relations $C \subset X \times Y$ and $C' \in Y \times Z$ is the relation

$$C \circ C' = \{(x, z) \in X \times Z : \text{ there exists } y \in Y$$
$$\text{such that } (x, y) \in C, (y, z) \in C'\}.$$

- The transposition of a relation C is the relation

$$C^T = \{(y, x) \in Y \times X : (x, y) \in C\}.$$

We call identity relation on X the set $I = \{(x, x) : x \in X\}$. Note that all our relations are "canonical" in the sense that they preserve areas in a generalized sense[2]. We will not purse this topic further here.

The high-level idea is that we would like to write $WF(Fm_1) \subset WF'(K) \circ WF(m_1)$, but this identity is not always true. The obstruction is that elements of $WF'(K)$ are in $T^*\mathbb{R}^{n \times n_d} \backslash 0$, which is in general a larger set than $(T^*\mathbb{R}^n \backslash 0) \times (T^*\mathbb{R}^{n_d} \backslash 0)$. Hence the interpretation of $WF'(K)$ as a relation runs into trouble with those elements such that either $\xi = 0$ but $(\xi_s, \xi_r, \omega) \neq 0$; or conversely $(\xi_s, \xi_r, \omega) = 0$ but $\xi \neq 0$. We say such elements are part of the *zero sections* of $WF'(K)$. In the sequel, we simply treat the case when those zero sections are empty.

The proper way of composing wavefront sets is the following celebrated theorem, due to Hörmander.

Theorem 5. *Let K be the distributional kernel of an operator F, with wavefront set (8.1). Assume that the elements of $WF'(K)$ obey*

$$\xi = 0 \quad \Leftrightarrow \quad (\xi_s, \xi_r, \omega) = 0,$$

i.e., $WF'(K)$ does not have zero sections. Then, for all m_1 in the domain of F,

$$WF(Fm_1) \subset WF'(K) \circ WF(m_1).$$

The physical significance of the zero sections will be made clear below when we compute $WF(K)$ for the imaging problem.

We can now consider the microlocal properties of the imaging operator F^*. Its kernel is simply K^T, the transpose of the kernel K of F. In turn, the relation $WF'(K^T)$ is simply the transpose of $WF'(K)$ seen as a relation in the product space $T^*\mathbb{R}^n \times T^*\mathbb{R}^{n_d}$,

$$WF'(K^T) = (WF'(K))^T.$$

[2] Namely, they are Lagrangian manifolds: the second fundamental symplectic form vanishes when restricted to canonical relations. The precaution of flipping the sign of the first covariable ξ in the definition of $WF'(K)$ translates into the fact that, in variables $((x, \xi), (y, \eta))$, it is $dx \wedge d\xi - dy \wedge d\eta$ that vanishes when restricted to the relation, and not $dx \wedge d\xi + dy \wedge d\eta$.

Our interest is in whether the normal operator F^*F preserves the singularities of the function m_1 it is applied to, i.e., whether imaging "places" singularities of the model perturbation at the right location from the knowledge of the linearized data Fm_1. To study the mapping of singularities of the normal operator F^*F, we are led to the composition

$$WF(F^*Fm_1) \subset (WF'(K))^T \circ WF'(K) \circ WF(m_1).$$

up to the same precaution involving zero-sections as above. The question is now whether the transpose relation can be considered an inverse, i.e., whether $(WF'(K))^T \circ WF'(K)$ is a subset of the identity on $T^*\mathbb{R}^n$. A simple condition of injectivity is necessary and sufficient for this to be the case.

Definition 7. *We say a relation $C \subset X \times Y$ is injective if*

$$(x_1, y) \in C, (x_2, y) \in C \quad \Rightarrow \quad x_1 = x_2.$$

As a map from X to Y, an injective relation may be multivalued, but as a map from Y to X, the transpose relation is a graph.

Lemma 3. *A relation C is injective if and only if $C^T \circ C \subset I$.*

Proof. Assume that $C \subset X \times Y$ is injective, and let $S \subset X$. By contradiction, if $C^T \circ C$ were not a subset of I, then there would exist an element x of X, for which there exists an element x' of $C^T \circ C \circ x$ such that $x' \neq x$. By definition of $C^T \circ C$, this means that there exists $y \in Y$ such that $(y, x') \in C^T$ and $(x, y) \in C$. By definition of transpose relation, we have $(x', y) \in C$. Injectivity of C implies $x = x'$, a contradiction.

By contraposition, assume that $C \subset X \times Y$ is not injective, i.e., there exist two elements (x, y) and (x', y) of C for which $x \neq x'$. Then $(y, x') \in C^T$ by definition of transpose, and $(x, x') \in C^T \circ C$ by definition of composition. Since $x \neq x'$, $C^T \circ C$ is not contained in the identity relation $I = \{(x, x) : x \in X\}$. □

We have all the pieces to gather the main result on preservation of singularities.

Theorem 6. *Let K be the distributional kernel of an operator F, with wavefront set (8.1). Assume that the elements of $WF'(K)$ obey*

$$\xi = 0 \quad \Leftrightarrow \quad (\xi_s, \xi_r, \omega) = 0,$$

i.e., $WF(K)$ does not have zero sections. Further assume that $WF'(K)$ is injective as a relation in $(T^*\mathbb{R}^n\backslash 0) \times (T^*\mathbb{R}^{n_d}\backslash 0)$. Then, for all m_1 in the domain of F,

$$WF(F^* F m_1) \subset WF(m_1).$$

Proof. Notice that the zero sections of $(WF'(K))^T$ are the same as those of $WF'(K)$, hence Hörmander's theorem can be applied twice. We get

$$WF(F^* F m_1) \subset (WF'(K))^T \circ WF'(K) \circ WF(m_1).$$

Under the injectivity assumption, lemma 3 allows to conclude. □

The assumptions of the theorem are tight:

- If zero sections are present, the composition law is different and adds elements away from the diagonal, as shown in the general formulation of Hörmander's theorem in [**?**].
- If injectivity does not hold, lemma 3 shows that the composition of two relations $C^T \circ C$ must have non-diagonal components.

8.2 Characterization of the wavefront set

In this section we construct the wavefront relation $WF'(K)$ explicitly in the simple scenario of section 7.2.

Recall that the setting of section 7.2 is one in which the background model $m_0(x)$ is heterogeneous and smooth, and enjoys no multipathing in the zone of interest. The source wavelet $w(t)$ (entering the right-hand side (!) of the wave equation) still needs to be an impulse $\delta(t)$ in order to create (!) propagating singularities. If instead the wavelet is essentially bandlimited, then so will the wavefields, but there are still remnants of the microlocal theory in the magnitude and locality of the propagating "wiggles"[3].

The distributional kernel of GRT migration is then

$$K(x; x_r, x_s, t) = a(x, x_r, x_s)\delta''(t - \tau(x_r, x, x_s)),$$

where $\tau(x_r, x, x_s) = \tau(x_r, x) + \tau(x, x_s)$ is the three-point traveltime, and a is a smooth amplitude, except for a singularity at $x = x_r$ and $x = x_s$.

[3]Microlocal theory can be expressed scale-by-scale, with wave packet decompositions and/or the so-called FBI transform. We leave out this topic.

The singular support of K is the same as its support as a measure, namely

$$\text{sing supp}(K) = \{(x, x_r, x_s, t) : t = \tau(x_r, x, x_s)\}.$$

In what follows we keep in mind that, for (x, x_r, x_s, t) to be a candidate base point in the wavefront relation of K, i.e., with a non-empty fiber, it is necessary that $t = \tau(x_r, x, x_s)$.

To find this wavefront relation, we first localize K around some reference point $(x_0, x_{r,0}, x_{s,0}, t_0)$ by multiplication with increasingly sharper cutoff functions, such as

$$\chi(\|x - x_0\|)\, \chi(\|x_r - x_{r,0}\|)\, \chi(\|x_s - x_{s,0}\|)\, \chi(|t - t_0|),$$

where χ is a C_0^∞ function compactly supported in the ball $B_\epsilon(0)$ for some ϵ that tends to zero. To keep notations manageable in the sequel, we introduce the symbol $[\,\chi\,]$ to refer to *any* C_0^∞ function of (x, x_r, x_s, t) (or any subset of those variables) with support in the ball of radius ϵ centered at $(x_0, x_{r,0}, x_{s,0}, t_0)$ (or any subset of those variables, resp.).

We then take a Fourier transform in every variable, and let

$$I(\xi, \xi_r, \xi_s, \omega) = \iiiint e^{i(x\cdot\xi - x_r\cdot\xi_r - x_s\cdot\xi_s - \omega t)} K(x; x_r, x_s, t)[\,\chi\,]dx\, dx_r\, dx_s\, dt. \tag{8.2}$$

According to the definition in the previous section, the test of membership in $WF'(K)$ involves the behavior of I under rescaling $(\xi, \xi_r, \xi_s, \omega) \to (\alpha\xi, \alpha\xi_r, \alpha\xi_s, \alpha\omega)$ by a single number $\alpha > 0$. Namely,

$$((x_0; \xi), (x_{r,0}, x_{s,0}, t_0; \xi_r, \xi_s, \omega)) \notin WF'(K)$$

provided there exists a sufficiently small $\epsilon > 0$ (determining the support of $[\,\chi\,]$) such that

$$I(\alpha\xi, \alpha\xi_r, \alpha\xi_s, \alpha\omega) \le C_m \alpha^{-m} \quad \text{for all } m > 0.$$

Notice that the Fourier transform in (8.2) is taken with a minus sign in the ξ variable (hence $e^{ix\cdot\xi}$ instead of $e^{-ix\cdot\xi}$) because $WF'(K)$ – the one useful for relations – precisely differs from $WF(K)$ by a minus sign in the ξ variable.

Before we simplify the quantity in (8.2), notice that it has a wave packet interpretation: we may let

$$\varphi(x) = e^{-ix\cdot\xi}\chi(\|x - x_0\|),$$

$$\psi(x_r, x_s.t) = e^{-i(x_r\cdot\xi_r + x_s\cdot\xi_s + \omega t)}\chi(\|x_r - x_{r,0}\|)\, \chi(\|x_s - x_{s,0}\|)\, \chi(|t - t_0|),$$

and view

$$I(\xi, \xi_r, \xi_s, \omega) = \langle \psi, F\varphi\rangle_{(x_r,x_s,t)},$$

with the complex conjugate over the second argument. The quantity $I(\xi, \xi_r, \xi_s, \omega)$ will be "large" (slow decay in α when the argument is scaled as earlier) provided the wave packet ψ in data space "matches", both in terms of location and oscillation content, the image of the wave packet φ in model space under the forward map F.

The t integral can be handled by performing two integrations by parts, namely

$$\int e^{-i\omega t}\delta''(t - \tau(x, x_r, x_s))\chi(\|t - t_0\|)\, dt = e^{-i\omega\tau(x,x_r,x_s)}\widetilde{\chi}_\omega(\tau(x, x_r, x_s)),$$

for some smooth function $\widetilde{\chi}_\omega(t) = e^{i\omega t}(e^{-i\omega t}\chi(|t-t_0|))''$ involving a (harmless) dependence on ω^0, ω^1 and ω^2. After absorbing the $\widetilde{\chi}_\omega(\tau(x, x_r, x_s))$ factor in a new cutoff function of (x, x_r, x_s) still called $[\,\chi\,]$, the result is

$$I(\xi, \xi_r, \xi_s, \omega) = \iiint e^{i(x\cdot\xi - x_r\cdot\xi_r - x_s\cdot\xi_s - \omega\tau(x,x_r,x_s))}a(x; x_r, x_s)[\,\chi\,]dx\, dx_r\, dx_s. \tag{8.3}$$

First, consider the case when $x_0 = x_{r,0}$ or $x_0 = x_{s,0}$. In that case, no matter how small $\epsilon > 0$, the cutoff function $[\,\chi\,]$ never avoids the singularity of the amplitude. As a result, the Fourier transform is expected to decay slowly in some directions. The ampltude's singularity is not the most interesting from a geometrical viewpoint, so for the sake of simplicity, we just brace for every possible $(\xi, \xi_r, \xi_s, \omega)$ to be part of the fiber relative $(x_0, x_{r,0}, x_{s,0}, t_0)$ where either $x_0 = x_{r,0}$ or $x_0 = x_{s,0}$. It is a good exercise to characterize these fibers in more details; see for instance the analysis in [**?**].

Assume now that $x_0 \neq x_{r,0}$ and $x_0 \neq x_{s,0}$. There exists a sufficiently small ϵ for which $a(x, x_r, x_s)$ is nonsingular and smooth on the support of $[\,\chi\,]$. We may therefore remove a from (8.3) by absorbing it in $[\,\chi\,]$:

$$I(\xi, \xi_r, \xi_s, \omega) = \iiint e^{i(x\cdot\xi - x_r\cdot\xi_r - x_s\cdot\xi_s - \omega\tau(x,x_r,x_s))}[\,\chi\,]dx\, dx_r\, dx_s. \tag{8.4}$$

What is left is an integral that can be estimated by the *stationary phase* lemma, or more precisely, the simplest version of such a result when the phase is nonstationary: see lemma 4 in appendix C.

The phase in (8.4) is $\phi(x, x_r, x_s) = x\cdot\xi - x_r\cdot\xi_r - x_s\cdot\xi_s - \omega\tau(x, x_r, x_s)$. Notice that ϕ involves the co-variables in a linear manner, hence is homogenenous of degree 1 in α, as needed in lemma 4. Its gradients are

$$\nabla_x \phi = \xi - \omega \nabla_x \tau(x, x_r, x_s),$$

$$\nabla_{x_r} \phi = -\xi_r - \omega \nabla_{x_r} \tau(x, x_r, x_s),$$

$$\nabla_{x_s} \phi = -\xi_s - \omega \nabla_{x_s} \tau(x, x_r, x_s).$$

If either of these gradients is nonzero at $(x_0, x_{r,0}, x_{s,0})$, then it will also be zero in a small neighborhood of that point, i.e., over the support of $[\chi]$ for ϵ small enough. In that case, lemma 4 applies, and it follows that the decay of the Fourier transform is fast no matter the (nonzero) direction $(\xi, \xi_r, \xi_s, \omega)$. Hence, if either of the gradients is nonzero, the point $((x_0, \xi), (x_{r,0}, x_{s,0}, t_0; \xi_r, \xi_s, \omega))$ is *not* in $WF'(K)$.

Hence $((x_0, \xi), (x_{r,0}, x_{s,0}, t_0; \xi_r, \xi_s, \omega))$ may be an element in $WF'(K)$ only if the phase has a critical point at $(x_0, x_{r,0}, x_{s,0})$:

$$\xi = \omega \nabla_x \tau(x_0, x_{r,0}, x_{s,0}), \tag{8.5}$$

$$\xi_r = -\omega \nabla_{x_r} \tau(x_0, x_{r,0}, x_{s,0}), \tag{8.6}$$

$$\xi_s = -\omega \nabla_{x_s} \tau(x_0, x_{r,0}, x_{s,0}). \tag{8.7}$$

Additionally, recall that $t_0 = \tau(x_0, x_{r,0}, x_{s,0})$. We gather the result as follows:

$$WF'(K) \subset \{ \, ((x, \xi), (x_r, x_s, t; \xi_r, \xi_s, \omega)) \in (T^*\mathbb{R}^n \times T^*\mathbb{R}^{n_d})\backslash 0 :$$

$$t = \tau(x, x_r, x_s), \ \omega \neq 0, \text{ and, either } x = x_r, \text{ or } x = x_s,$$

$$\text{or (8.5, 8.6, 8.7) hold at } (x, x_r, x_s) \, \}.$$

Whether the inclusion is a proper inclusion, or an equality, (at least away from $x = x_r$ and $x = x_s$) depends on whether the amplitude factor in (8.4) vanishes in an open set or not.

Notice that $\omega \neq 0$ for the elements in $WF'(K)$, otherwise (8.5, 8.6, 8.7) would imply that the other covariables are zero as well, which is not allowed in the definition of $WF'(K)$. In the sequel, we may then divide by ω at will.

The relations (8.5, 8.6, 8.7) have an important physical meaning. Recall that $t = \tau(x, x_r, x_s)$ is called isochrone curve/surface when it is considered in model x-space, and moveout curve/surface when considered in data (x_r, x_s, t)-space.

- The relation $\xi = \omega \nabla_x \tau(x, x_r, x_s)$ indicates that ξ is *normal to the isochrone* passing through x, with level set $t = \tau(x, x_r, x_s)$. In terms of two-point traveltimes, we may write

$$\frac{\xi}{\omega} = \nabla_x \tau(x, x_r) + \nabla_x \tau(x, x_s).$$

  Observe that $\nabla_x \tau(x, x_r)$ is tangent at x to the ray from x_r to x, and $\nabla_x \tau(x, x_s)$ is tangent at x to the ray from x_s to x, hence ξ is the bisector direction for those two rays. The (co)vector ξ may be understood as the (co)normal to a "localized mirror" about which the incident wave reflects in a specular manner to create the scattered wave. The equation above is then the vector expression of *Snell's law of reflection*, that the angle of incidence $\angle(\nabla_x \tau(x, x_s), \xi)$ equals the angle of reflection $\angle(\nabla_x \tau(x, x_r), \xi)$.

- The special case $\xi = 0$ is equally important from a physical viewpoint. Since $\omega \neq 0$, it corresponds to

$$\nabla_x \tau(x, x_r) = -\nabla_x \tau(x, x_s),$$

  i.e., the tangents to the incident and reflected rays are collinear and opposite in signs. This happens when x is a point on the direct, unbroken ray linking x_r to x_s. We may call this situation forward scattering: it corresponds to *transmitted waves* rather than reflected waves. The reader can check that it is the only way in which a zero section is created in the wavefront relation (see the previous section for an explanation of zero sections and how they impede the interpretation of $WF'(K)$ as a relation.)

- Consider now (8.6) and (8.7). Pick any $\eta = (\eta_r, \eta_s)$, form the combination $\eta_r \cdot$ (8.6) + $\eta_s \cdot$ (8.7) = 0, and rearrange the terms to get

$$\begin{pmatrix} \xi_r \\ \xi_s \\ \omega \end{pmatrix} \cdot \begin{pmatrix} \eta_r \\ \eta_s \\ \eta_r \cdot \nabla_{x_r} \tau + \eta_s \cdot \nabla_{x_s} \tau \end{pmatrix} = 0,$$

  with both τ evaluated at (x, x_r, x_s). The second vector in the dot product is an arbitrary vector tangent to the moveout surface $t = \tau(x, x_r, x_s)$ in (x_r, x_s, t)-data space. Thus (ξ_r, ξ_s, ω) is *normal to the moveout surface.*

(!) We are now ready to interpret $WF'(K)$ in terms of the mapping of singularities that it generates. Assume that no forward scattering is taking place, i.e., we are only dealing with reflected rather than transmitted waves. From a singularity in model space at a point x and in the direction ξ – a local mirror placed at x with conormal ξ – the wavefront relation predicts that there may be singularities in data space, with location(s) (x_r, x_s, t) and corresponding conormal direction(s) (ξ_r, ξ_s, ω) determined as follows.

Fix a couple (x, ξ).

1. For each x_s, find the unique ray that links x_s to x.

2. At x, find the direction of the incoming ray as $\nabla_x \tau(x, x_s)$.

3. Determine $\nabla_x \tau(x, x_r)$ as the vector with magnitude $1/c(x)$ (from the eikonal equation), and direction so that (8.5) holds for some ω. This can be done by computing the reflection of $\nabla_x \tau(x, x_s)$ about the axis generated by ξ,

$$\nabla_x \tau(x, x_r) = \nabla_x \tau(x, x_s) - 2\frac{\nabla_x \tau(x, x_s) \cdot \xi}{\xi \cdot \xi} \xi.$$

4. Trace the ray from x and take-off direction $-\nabla_x \tau(x, x_r)$:

   - If the ray is closed or exits the domain before reaching a receiver, discard x_s. This particular x_s does not give rise to any singularity (relative to this particular couple (x, ξ).)

   - If the ray reaches a receiver x_r, let $t = \tau(x, x_s) + \tau(x, x_r)$. Some singularity may appear at the point (x_s, x_r, t) in data space.

5. Determine $\omega = \|\xi\| / \|\nabla_x \tau(x, x_r) + \nabla_x \tau(x, x_s)\|$. The normalization of ξ starts as arbitrary, and is undone by division by ω.

6. Determine ξ_r and ξ_s directly from (8.6) and (8.7). Then the covector conormal to the singularity at (x_r, x_s, t) is (ξ_r, ξ_s, ω) (as well as all its positive multiples.)

Notice that both the data variables (x_r, x_s, t) and the covariables (ξ_r, ξ_s, ω) are determined uniquely from x and ξ in this situation, hence the singularity mapping induced by the wavefront relation is one-to-one.

If we make the additional two assumptions that the receivers surround the domain of interest (which would require us to give up the assumption that the acquisition manifold Ω is a vector subspace), and that there is no trapped (closed) ray , then the test in point 4 above always succeeds. In that case, the wavefront relation is onto, hence bijective, and can be inverted by a sequence of steps with a similar geometrical content as earlier. See an exercise at the end of this chapter. (!)

8.3 Pseudodifferential theory

8.4 Exercises

1. Compute the wavefront set of the following functions/distributions of two variables: (1) the indicator function $H(x_1)$ (where H is the Heaviside function) of the half-plane $x_1 > 0$, (2) the indicator function of the unit disk, (3) the indicator function of the square $[-1, 1]^2$, (4) the Dirac delta $\delta(x)$.

2. Find conditions on the acquisition manifold under which forward scattering is the only situation in which a zero section can be created in $WF'(K)$. [Hint: transversality with the rays from x to x_r, and from x to x_s.]

3. Perform the geometrical construction of $(WF'(K))^T$ for migration, analogous to the geometrical construction of $WF'(K)$ for Born modeling done in section 8.2. Assume: 1) τ single-valued, 2) no forward scattering, 3) full aperture of receivers, 4) no trapped rays, and 5) the acquisition manifold Ω is a vector subspace equal to the Cartesian product of two copies of the same subspace of codimension 1 in $\mathbb{R}^3$ (one for x_r and one for x_s). In other words, you may assume that both x_r and x_s lie on the "surface" $z = 0$.

   Solution. Fix (x_r, x_s, t) and $(\xi_r, \xi_s, \omega) \neq 0$. We aim to determine a unique (x, ξ) so that $((x, \xi), (x_r, x_s, t; \xi_r, \xi_s, \omega)) \in WF'(K)$.

   Starting with (8.6), notice first that $\nabla_{x_r} \tau(x, x_r)$ is a partial gradient: it is the projection to the acquisition manifold Ω at x_r of the three-dimensional gradient of τ in its second argument, say $\nabla_y \tau(x, y)|_{y=(x_r,0)}$, in coordinates $y = (x_r, z)$ so that Ω is represented by $z = 0$. A similar observation holds for $\nabla_{x_s}\tau(x, x_s)$. Since the norm of each partial gradient is less than the norm of the full gradient, and since the eikonal equation determines this latter norm, ξ_r and ξ_s must obey the geometric compatibility relations

$$|\xi_r| \leq \frac{\omega}{c(x_r)}, \qquad |\xi_s| \leq \frac{\omega}{c(x_s)}.$$

The coordinate z is one-dimensional from our assumption on the acquisition manifold, so we can now assemble the full gradient of τ at x_r. With the help of (8.6) we obtain

$$\nabla_y \tau(x, y)|_{y=(x_r,0)} = \left(-\frac{\xi_r}{\omega}, \pm\sqrt{\frac{1}{c^2(x_r)} - \frac{|\xi_r|^2}{\omega^2}} \right).$$

(The sign is determined on geometrical grounds, so that $\nabla_y \tau$ points outside the domain of interest.) Minus the direction of $\nabla_y \tau$ is also the take-off direction of a ray from x_r. We obtain $\nabla_y \tau(x, y)|_{y=(x_s,0)}$ in the same fashion from (8.7), which determines the take-off direction of a ray from x_s. These rays meet at a single point that we call x. We may alternatively intersect either of these rays with the isochrone $t = \tau(x, x_r, x_s)$ to obtain x; this piece of information is redundant in the wavefront relation[4]. Finally, we determine $\xi = \omega \nabla_x \tau(x, x_r, x_s)$.

[4] Only because we have assumed that the acquisition manifold has low codimension.