Josselin Garnier

Inverse Problems and Imaging

January 6, 2024

Contents

1

The Green's function of the wave equation

In this chapter we review a few important properties of the Green's function, the fundamental solution of the wave equation. These properties -reciprocity (Proposition 1.1) and Helmholtz-Kirchhoff identity (Proposition 1.2)- are instrumental in the analysis of time-reversal experiments and imaging.

We consider the scalar wave model:

$$\frac{1}{c^2(\boldsymbol{x})}\frac{\partial^2 u}{\partial t^2} - \Delta_{\boldsymbol{x}} u = n(t,\boldsymbol{x}). \tag{1.1}$$

Here $n(t,\boldsymbol{x})$ is the source term and $c(\boldsymbol{x})$ is the propagation speed (parameter of the medium), which is assumed to be constant outside a domain with compact support. We consider the three-dimensional case $\boldsymbol{x} \in \mathbb{R}^3$.

1.1 Green's function

We introduce the time-dependent Green's function $G(t,\boldsymbol{x},\boldsymbol{y})$ which is the fundamental solution of

$$\frac{1}{c^2(\boldsymbol{x})}\frac{\partial^2 G}{\partial t^2} - \Delta_{\boldsymbol{x}} G = \delta(t)\delta(\boldsymbol{x}-\boldsymbol{y}), \tag{1.2}$$

with the initial conditions

$$G(t,\boldsymbol{x},\boldsymbol{y}) = 0, \qquad \forall t < 0. \tag{1.3}$$

It corresponds to the wave propagating in the medium when there is a point source at $\boldsymbol{y}$ emitting a Dirac pulse at time 0.

If the medium is homogeneous $c(\boldsymbol{x}) \equiv c_0$, then

$$G(t,\boldsymbol{x},\boldsymbol{y}) = \frac{1}{4\pi|\boldsymbol{x}-\boldsymbol{y}|}\delta\Big(t - \frac{|\boldsymbol{x}-\boldsymbol{y}|}{c_0}\Big), \qquad t > 0,$$

which corresponds to a spherical wave propagating at speed c_0 (see for instance [22, Chapter 2] or [20]).

The time-harmonic Green's function is the Fourier transform of the time-dependent Green's function:

$$\hat{G}(\omega, \boldsymbol{x}, \boldsymbol{y}) = \int G(t, \boldsymbol{x}, \boldsymbol{y}) e^{i\omega t} \mathrm{d}t.$$

It is the solution of the Helmholtz equation

$$\Delta_{\boldsymbol{x}} \hat{G} + \frac{\omega^2}{c^2(\boldsymbol{x})} \hat{G} = -\delta(\boldsymbol{x} - \boldsymbol{y}), \tag{1.4}$$

and it satisfies the Sommerfeld radiation condition ($c(\boldsymbol{x}) = c_0$ at infinity):

$$\lim_{|\boldsymbol{x}| \to \infty} |\boldsymbol{x}| \Big(\frac{\boldsymbol{x}}{|\boldsymbol{x}|} \cdot \nabla_{\boldsymbol{x}} - i \frac{\omega}{c_0} \Big) \hat{G}(\omega, \boldsymbol{x}, \boldsymbol{y}) = 0. \tag{1.5}$$

Moreover, the convergence is uniform with respect to $\boldsymbol{x}/|\boldsymbol{x}|$ over the sphere and the Sommerfeld radiation condition also implies that

$$\limsup_{|\boldsymbol{x}| \to \infty} |\boldsymbol{x}| |\hat{G}(\omega, \boldsymbol{x}, \boldsymbol{y})| \tag{1.6}$$

is bounded uniformly with respect to $\boldsymbol{x}/|\boldsymbol{x}|$ over the sphere. The time-harmonic Green's function corresponds to the wave propagating in the medium when there is a point source at $\boldsymbol{y}$ emitting a time-harmonic signal at frequency ω.

If the medium is homogeneous $c(\boldsymbol{x}) \equiv c_0$, then

$$\hat{G}(\omega, \boldsymbol{x}, \boldsymbol{y}) = \frac{1}{4\pi |\boldsymbol{x} - \boldsymbol{y}|} e^{i \frac{\omega}{c_0} |\boldsymbol{x} - \boldsymbol{y}|}. \tag{1.7}$$

The solution of the wave equation with source $n(t, \boldsymbol{x})$ can be expressed as the convolution of the Green's function and the source term:

$$u(t, \boldsymbol{x}) = \iint G(t - s, \boldsymbol{x}, \boldsymbol{y}) n(s, \boldsymbol{y}) \mathrm{d}\boldsymbol{y} \mathrm{d}s. \tag{1.8}$$

Throughout these notes, the integrals with respect to a time variable are carried over $(-\infty, \infty)$ and the integrals with respect to a space variable are carried over $\mathbb{R}^3$, unless stated otherwise. Note however that, by (1.3), the integral (1.8) is causal:

$$u(t, \boldsymbol{x}) = \int_{-\infty}^{t} \int_{\mathbb{R}^3} G(t - s, \boldsymbol{x}, \boldsymbol{y}) n(s, \boldsymbol{y}) \mathrm{d}\boldsymbol{y} \mathrm{d}s$$

in the sense that u at time t only depends on the values of the source term for times smaller than t.

The Fourier transform of the solution

$$\hat{u}(\omega, \boldsymbol{x}) = \int u(t, \boldsymbol{x}) e^{i\omega t} \mathrm{d}t$$

is simply given by:

$$\hat{u}(\omega, \boldsymbol{x}) = \int \hat{G}(\omega, \boldsymbol{x}, \boldsymbol{y}) \hat{n}(\omega, \boldsymbol{y}) \mathrm{d}\boldsymbol{y}.$$

1.2 Sommerfeld radiation condition

The Sommerfeld radiation condition plays an important role. A solution of the Helmholtz equation is called radiating if it satisfies the Sommerfeld radiation condition (1.5) uniformly in all directions ($c(\boldsymbol{x}) = c_0$ at infinity).

The Helmholtz equation (1.4) has an infinite number of solutions. For example, if $c(\boldsymbol{x}) \equiv c_0$, then, for any a, the function

$$\hat{G}_a(\omega, \boldsymbol{x}, \boldsymbol{y}) = \frac{1-a}{4\pi|\boldsymbol{x}-\boldsymbol{y}|} \exp\left(i\frac{\omega}{c_0}|\boldsymbol{x}-\boldsymbol{y}|\right) + \frac{a}{4\pi|\boldsymbol{x}-\boldsymbol{y}|} \exp\left(-i\frac{\omega}{c_0}|\boldsymbol{x}-\boldsymbol{y}|\right)$$

is solution of the homogeneous Helmholtz equation. However, only the solution with $a = 0$ satisfies the Sommerfeld radiation condition. It corresponds to a field radiating from $\boldsymbol{y}$. The other solutions are "unphysical". For example, the solution with $a = 1$ can be interpreted as energy coming from infinity and sinking at $\boldsymbol{y}$.

The important result is the following theorem (see [17, Vol. 2, Chap. IV, Sec. 5] for a classical statement or [35] for a more detailed presentation): The Helmholtz equation (with c bounded and constant outside a compact) has a unique radiating solution.

1.3 Reciprocity

An important property satisfied by the Green's function is the reciprocity property:

Proposition 1.1. *For any $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^3$, we have*

$$\hat{G}(\omega, \boldsymbol{x}, \boldsymbol{y}) = \hat{G}(\omega, \boldsymbol{y}, \boldsymbol{x}). \tag{1.9}$$

This result means that the wave recorded at $\boldsymbol{x}$ when there is a time-harmonic source at $\boldsymbol{y}$ is equal to the wave recorded at $\boldsymbol{y}$ when there is a time-harmonic source at $\boldsymbol{x}$. It is obvious in the homogeneous medium case from the explicit expression (1.7) of the time-harmonic Green's function. In the general case it follows from the divergence theorem as shown in the following proof. This

proof is standard but we give it because its key arguments will be used several times in the forthcoming sections.

Proof. We consider the equations satisfied by the Green's function with the source at $\boldsymbol{y}_2$ and with the source at $\boldsymbol{y}_1$ (with $\boldsymbol{y}_1 \neq \boldsymbol{y}_2$):

$$\Delta_{\boldsymbol{x}}\hat{G}(\omega,\boldsymbol{x},\boldsymbol{y}_2)+\frac{\omega^2}{c^2(\boldsymbol{x})}\hat{G}(\omega,\boldsymbol{x},\boldsymbol{y}_2)=-\delta(\boldsymbol{x}-\boldsymbol{y}_2),$$
$$\Delta_{\boldsymbol{x}}\hat{G}(\omega,\boldsymbol{x},\boldsymbol{y}_1)+\frac{\omega^2}{c^2(\boldsymbol{x})}\hat{G}(\omega,\boldsymbol{x},\boldsymbol{y}_1)=-\delta(\boldsymbol{x}-\boldsymbol{y}_1).$$

We multiply the first equation by $\hat{G}(\omega,\boldsymbol{x},\boldsymbol{y}_1)$ and subtract the second equation multiplied by $\hat{G}(\omega,\boldsymbol{x},\boldsymbol{y}_2)$:

$$\begin{aligned}&\nabla_{\boldsymbol{x}}\cdot\Big[\hat{G}(\omega,\boldsymbol{x},\boldsymbol{y}_1)\nabla_{\boldsymbol{x}}\hat{G}(\omega,\boldsymbol{x},\boldsymbol{y}_2)-\hat{G}(\omega,\boldsymbol{x},\boldsymbol{y}_2)\nabla_{\boldsymbol{x}}\hat{G}(\omega,\boldsymbol{x},\boldsymbol{y}_1)\Big]\\&=\hat{G}(\omega,\boldsymbol{x},\boldsymbol{y}_2)\delta(\boldsymbol{x}-\boldsymbol{y}_1)-\hat{G}(\omega,\boldsymbol{x},\boldsymbol{y}_1)\delta(\boldsymbol{x}-\boldsymbol{y}_2)\\&=\hat{G}(\omega,\boldsymbol{y}_1,\boldsymbol{y}_2)\delta(\boldsymbol{x}-\boldsymbol{y}_1)-\hat{G}(\omega,\boldsymbol{y}_2,\boldsymbol{y}_1)\delta(\boldsymbol{x}-\boldsymbol{y}_2).\end{aligned}$$

We next integrate over the ball $B(\mathbf{0},L)$ with center at $\mathbf{0}$ and radius L which contains both $\boldsymbol{y}_1$ and $\boldsymbol{y}_2$ and use the divergence theorem:

$$\begin{aligned}&\int_{\partial B(\mathbf{0},L)}\boldsymbol{n}(\boldsymbol{x})\cdot\Big[\hat{G}(\omega,\boldsymbol{x},\boldsymbol{y}_1)\nabla_{\boldsymbol{x}}\hat{G}(\omega,\boldsymbol{x},\boldsymbol{y}_2)-\hat{G}(\omega,\boldsymbol{x},\boldsymbol{y}_2)\nabla_{\boldsymbol{x}}\hat{G}(\omega,\boldsymbol{x},\boldsymbol{y}_1)\Big]\mathrm{d}\sigma(\boldsymbol{x})\\&=\hat{G}(\omega,\boldsymbol{y}_1,\boldsymbol{y}_2)-\hat{G}(\omega,\boldsymbol{y}_2,\boldsymbol{y}_1),\end{aligned}$$

where $\boldsymbol{n}(\boldsymbol{x})$ is the unit outward normal to the ball $B(\mathbf{0},L)$, which is $\boldsymbol{n}(\boldsymbol{x})=\boldsymbol{x}/|\boldsymbol{x}|$.

If $\boldsymbol{x}\in\partial B(\mathbf{0},L)$ and $L\to\infty$, then we have by the Sommerfeld radiation condition:

$$\boldsymbol{n}\cdot\nabla_{\boldsymbol{x}}\hat{G}(\omega,\boldsymbol{x},\boldsymbol{y})=i\frac{\omega}{c_0}\hat{G}(\omega,\boldsymbol{x},\boldsymbol{y})+o\Big(\frac{1}{L}\Big).$$

Since we also have $\hat{G}(\omega,\boldsymbol{x},\boldsymbol{y})=O(1/L)$, we find that, as $L\to\infty$,

$$\begin{aligned}&\hat{G}(\omega,\boldsymbol{y}_1,\boldsymbol{y}_2)-\hat{G}(\omega,\boldsymbol{y}_2,\boldsymbol{y}_1)\\&=i\frac{\omega}{c_0}\int_{\partial B(\mathbf{0},L)}\hat{G}(\omega,\boldsymbol{x},\boldsymbol{y}_1)\hat{G}(\omega,\boldsymbol{x},\boldsymbol{y}_2)-\hat{G}(\omega,\boldsymbol{x},\boldsymbol{y}_2)\hat{G}(\omega,\boldsymbol{x},\boldsymbol{y}_1)\mathrm{d}\sigma(\boldsymbol{x})\\&=0,\end{aligned}$$

which is the desired result. □

1.4 Helmholtz-Kirchhoff identity

Let us consider the following set up: the observation points are $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$, the medium may be inhomogeneous is a region within the ball $B(\mathbf{0},D)$ with center at $\mathbf{0}$ and radius D (see Figure 1.1).

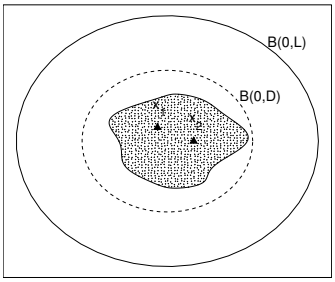


Fig. 1.1. Situation for the Helmholtz-Kirchhoff identity: The medium is homogeneous outside the ball $B(\mathbf{0}, D)$.

The Helmholtz-Kirchhoff identity (1.11) is stated in the second part of the following theorem. It follows from the divergence theorem and the Sommerfeld radiation condition. This identity is known in acoustics [8, p. 473] and in optics [11, p. 419]) and it is also presented in [37]. Note that (1.11) relates a product of two time-harmonic Green's functions, one of them being complex conjugated, to a time-harmonic Green's function. The product, in the time domain, has the form of a correlation. This explains the importance of the Helmholtz-Kirchhoff identity for correlation-based imaging as we will see later.

Theorem 1.2. *1) For any bounded and smooth open domain Ω, the second Green's identity holds for any $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \Omega$:*

$$\begin{aligned} &\hat{G}(\omega, \boldsymbol{x}_1, \boldsymbol{x}_2) - \overline{\hat{G}(\omega, \boldsymbol{x}_1, \boldsymbol{x}_2)} \qquad (1.10)\\ &= \int_{\partial\Omega} \boldsymbol{n} \cdot \left[\overline{\hat{G}(\omega, \boldsymbol{y}, \boldsymbol{x}_1)} \nabla_{\boldsymbol{y}} \hat{G}(\omega, \boldsymbol{y}, \boldsymbol{x}_2) - \hat{G}(\omega, \boldsymbol{y}, \boldsymbol{x}_2) \nabla_{\boldsymbol{y}} \overline{\hat{G}(\omega, \boldsymbol{y}, \boldsymbol{x}_1)}\right] \mathrm{d}\sigma(\boldsymbol{y}), \end{aligned}$$

where $\boldsymbol{n}$ is the unit outward normal to the Ω.

2) If the medium is homogeneous (velocity c_0) outside $B(\mathbf{0}, D)$, then $\forall \boldsymbol{x}_1, \boldsymbol{x}_2 \in B(\mathbf{0}, D)$ we have for $L \gg D$:

$$\hat{G}(\omega, \boldsymbol{x}_1, \boldsymbol{x}_2) - \overline{\hat{G}(\omega, \boldsymbol{x}_1, \boldsymbol{x}_2)} = \frac{2i\omega}{c_0} \int_{\partial B(\mathbf{0}, L)} \overline{\hat{G}(\omega, \boldsymbol{x}_1, \boldsymbol{y})} \hat{G}(\omega, \boldsymbol{x}_2, \boldsymbol{y}) \mathrm{d}\sigma(\boldsymbol{y}). \qquad (1.11)$$

The Helmholtz-Kirchhoff identity turns out to be very useful when we analyze the scattering of waves by an obstacle, the refocusing of waves during a time-reversal experiment, and the cross correlation of signals emitted by ambient noise sources, as we will see later.

Proof. The proof is based essentially on the divergence theorem and the Sommerfeld radiation condition, as the one of reciprocity that was given in the previous section. Let us consider

$$\Delta_{\boldsymbol{y}}\hat{G}(\omega,\boldsymbol{y},\boldsymbol{x}_2)+\frac{\omega^2}{c^2(\boldsymbol{y})}\hat{G}(\omega,\boldsymbol{y},\boldsymbol{x}_2)=-\delta(\boldsymbol{y}-\boldsymbol{x}_2),$$
$$\Delta_{\boldsymbol{y}}\overline{\hat{G}(\omega,\boldsymbol{y},\boldsymbol{x}_1)}+\frac{\omega^2}{c^2(\boldsymbol{y})}\overline{\hat{G}(\omega,\boldsymbol{y},\boldsymbol{x}_1)}=-\delta(\boldsymbol{y}-\boldsymbol{x}_1).$$

We multiply the first equation by $\overline{\hat{G}(\omega,\boldsymbol{y},\boldsymbol{x}_1)}$ and we subtract the second equation multiplied by $\hat{G}(\omega,\boldsymbol{y},\boldsymbol{x}_2)$:

$$\begin{aligned}
&\nabla_{\boldsymbol{y}}\cdot\Big[\overline{\hat{G}(\omega,\boldsymbol{y},\boldsymbol{x}_1)}\nabla_{\boldsymbol{y}}\hat{G}(\omega,\boldsymbol{y},\boldsymbol{x}_2)-\hat{G}(\omega,\boldsymbol{y},\boldsymbol{x}_2)\nabla_{\boldsymbol{y}}\overline{\hat{G}(\omega,\boldsymbol{y},\boldsymbol{x}_1)}\Big]\\
&=\hat{G}(\omega,\boldsymbol{y},\boldsymbol{x}_2)\delta(\boldsymbol{y}-\boldsymbol{x}_1)-\overline{\hat{G}(\omega,\boldsymbol{y},\boldsymbol{x}_1)}\delta(\boldsymbol{y}-\boldsymbol{x}_2)\\
&=\hat{G}(\omega,\boldsymbol{x}_1,\boldsymbol{x}_2)\delta(\boldsymbol{y}-\boldsymbol{x}_1)-\overline{\hat{G}(\omega,\boldsymbol{x}_1,\boldsymbol{x}_2)}\delta(\boldsymbol{y}-\boldsymbol{x}_2),
\end{aligned}$$

using the reciprocity property $\hat{G}(\omega,\boldsymbol{x}_2,\boldsymbol{x}_1)=\hat{G}(\omega,\boldsymbol{x}_1,\boldsymbol{x}_2)$.

We integrate over the domain Ω and we use the divergence theorem to obtain (1.10). When Ω is the ball $B(\boldsymbol{0},L)$, then the unit outward normal is $\boldsymbol{n}=\boldsymbol{y}/|\boldsymbol{y}|$. The Green's function also satisfies the Sommerfeld radiation condition

$$\lim_{|\boldsymbol{y}|\to\infty}|\boldsymbol{y}|\Big(\frac{\boldsymbol{y}}{|\boldsymbol{y}|}\cdot\nabla_{\boldsymbol{y}}-\ i\frac{\omega}{c_0}\Big)\hat{G}(\omega,\boldsymbol{y},\boldsymbol{x}_1)=0,$$

uniformly in all directions $\boldsymbol{y}/|\boldsymbol{y}|$. Using this property, we substitute

$$i\frac{\omega}{c_0}\hat{G}(\omega,\boldsymbol{y},\boldsymbol{x}_2)\ \text{for}\ \boldsymbol{n}\cdot\nabla_{\boldsymbol{y}}\hat{G}(\omega,\boldsymbol{y},\boldsymbol{x}_2)$$

in the surface integral over $\partial B(\boldsymbol{0},L)$, and

$$-i\frac{\omega}{c_0}\overline{\hat{G}(\omega,\boldsymbol{y},\boldsymbol{x}_1)}\ \text{for}\ \boldsymbol{n}\cdot\nabla_{\boldsymbol{y}}\overline{\hat{G}(\omega,\boldsymbol{y},\boldsymbol{x}_1)},$$

and we obtain the desired result. □

1.5 Application to time reversal

In this subsection we will apply the Helmholtz-Kirchhoff identity to study a time-reversal experiment. Originally time reversal was proposed not for imaging, but for energy focusing. The idea was to focus ultrasound energy on kidney stones to destroy them [21]. However the analysis of this situation has triggered most of the work on reverse-time migration and also correlation-based imaging as we will see later. That is why we first study this idea.

A time-reversal experiment is based on the use of a special device called time-reversal mirror, which is an array of transducers that can be used as sources or as receivers. In a time-reversal experiment the time-reversal mirror is used first as a receiver array, then as a source array. Let us consider the ideal

case in which the time-reversal mirror covers the surface of the ball $B(\mathbf{0}, L)$ (see Figure 1.2).

In the first step of a time-reversal experiment (Figure 1.2, left), a point source at $\boldsymbol{y}$ emits a pulse $f(t)$. The waves $\hat{u}(\omega, \boldsymbol{x})$ are recorded at the surface $\partial B(\mathbf{0}, L)$:

$$\hat{u}(\omega, \boldsymbol{x}) = \hat{G}(\omega, \boldsymbol{x}, \boldsymbol{y})\hat{f}(\omega), \quad \boldsymbol{x} \in \partial B(\mathbf{0}, L).$$

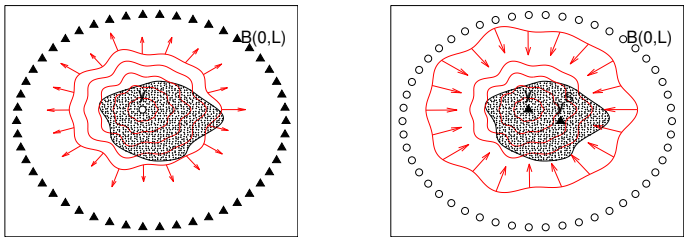


Fig. 1.2. A time-reversal experiment with full-aperture time-reversal mirror. In the first step (left), the time-reversal mirror is used as an array of receivers. In the second step (right), the time-reversal mirror is used as an array of sources.

In the second step of the experiment (Figure 1.2, right), the recorded signals are time-reversed and sent back into the medium. The signal received at $\boldsymbol{y}^S$ is $\hat{u}_{\mathrm{TR}}(\omega, \boldsymbol{y}^S)$:

$$\hat{u}_{\mathrm{TR}}(\omega, \boldsymbol{y}^S) = \int_{\partial B(\mathbf{0},L)} \mathrm{d}\sigma(\boldsymbol{x})\hat{G}(\omega, \boldsymbol{y}^S, \boldsymbol{x})\overline{\hat{G}(\omega, \boldsymbol{x}, \boldsymbol{y})\hat{f}(\omega)},$$

where we have used the fact that the Fourier transform of the time reverse of a real-valued function is the complex conjugate of the Fourier transform of the function. Using reciprocity $\hat{G}(\omega, \boldsymbol{x}, \boldsymbol{y}) = \hat{G}(\omega, \boldsymbol{y}, \boldsymbol{x})$:

$$\hat{u}_{\mathrm{TR}}(\omega, \boldsymbol{y}^S) = \int_{\partial B(\mathbf{0},L)} \mathrm{d}\sigma(\boldsymbol{x})\overline{\hat{G}(\omega, \boldsymbol{y}, \boldsymbol{x})}\hat{G}(\omega, \boldsymbol{y}^S, \boldsymbol{x})\overline{\hat{f}(\omega)}.$$

By Helmholtz-Kirchhoff identity:

$$\hat{G}(\omega, \boldsymbol{y}, \boldsymbol{y}^S) - \overline{\hat{G}(\omega, \boldsymbol{y}, \boldsymbol{y}^S)} = \frac{2i\omega}{c_0} \int_{\partial B(\mathbf{0},L)} \mathrm{d}\sigma(\boldsymbol{x})\overline{\hat{G}(\omega, \boldsymbol{y}, \boldsymbol{x})}\hat{G}(\omega, \boldsymbol{y}^S, \boldsymbol{x}),$$

we get

$$\hat{u}_{\mathrm{TR}}(\omega, \boldsymbol{y}^S) = \frac{\hat{G}(\omega, \boldsymbol{y}, \boldsymbol{y}^S) - \overline{\hat{G}(\omega, \boldsymbol{y}, \boldsymbol{y}^S)}}{2i\omega/c_0}\overline{\hat{f}(\omega)}.$$

Remember that $\boldsymbol{y}$ is the original source location. Therefore, the focal spot, i.e. the spatial form of the time-reversed wave, is determined by the imaginary part of the Green's function:

$$\hat{u}_{\mathrm{TR}}(\omega, \boldsymbol{y}^S) = \frac{c_0}{\omega} \mathrm{Im}\big(\hat{G}(\omega, \boldsymbol{y}, \boldsymbol{y}^S)\big)\overline{\hat{f}(\omega)}. \tag{1.12}$$

In a homogeneous medium with propagation speed c_0 we have

$$\hat{G}(\omega, \boldsymbol{y}, \boldsymbol{y}^S) = \frac{1}{4\pi|\boldsymbol{y} - \boldsymbol{y}^S|} e^{i\frac{\omega|\boldsymbol{y}-\boldsymbol{y}^S|}{c_0}},$$

which gives the focal spot

$$\hat{u}_{\mathrm{TR}}(\omega, \boldsymbol{y}^S) = \frac{1}{4\pi}\mathrm{sinc}\Big(\frac{\omega|\boldsymbol{y} - \boldsymbol{y}^S|}{c_0}\Big)\overline{\hat{f}(\omega)}, \qquad \mathrm{sinc}(s) := \frac{\sin s}{s}. \tag{1.13}$$

This shows that the focal spot has a diameter equal to $\lambda/2$, where $\lambda = 2\pi c_0/\omega$ is the wavelength. This is the so-called diffraction limit.

Remark: In a complex medium $\mathrm{Im}\big(\hat{G}(\omega, \boldsymbol{y}, \boldsymbol{y}^S)\big)$ can be sharper than in a homogeneous medium. This is the "super-resolution effect" mentioned in the recent literature [31]: if a micro-structured medium surrounds the original source $\boldsymbol{y}$, then the focal spot can be smaller than the diffraction limit $\lambda/2$. Indeed, the main effect of the micro-structured medium is to modify the effective wavelength. This can be proved by an asymptotic analysis (homogeneization result) [4, 28, 32].

In this subsection we have presented a proof that time reversal of waves results in a refocusing at the original source location. This proof is based on the Helmholtz-Kirchhoff identity, which requires that the time-reversal mirror completely surrounds the region of interest. This is the full-aperture case. However, time-reversal refocusing happens also with spatially limited time-reversal mirrors. This is the partial aperture case. This has been shown experimentally, numerically, and theoretically. The proof that time-reversal refocusing happens in the partial aperture case cannot use the Helmholtz-Kirchhoff identity, but rather follows from asymptotic theory for wave propagation (especially in the high-frequency regime). The size of the focal spot is then larger than the diffraction limit, and takes the form of the Rayleigh resolution formula $\theta\lambda$, where θ is the aperture angle, that is, the ratio of the array diameter over the distance from the array to the source.

1.6 The scalar wave equation with noise sources

We consider the solution u of the wave equation (1.1) in a three-dimensional inhomogeneous medium with propagation speed $c(\boldsymbol{x})$:

$$\frac{1}{c^2(\boldsymbol{x})}\frac{\partial^2 u}{\partial t^2} - \Delta_{\boldsymbol{x}} u = n(t, \boldsymbol{x}). \tag{1.14}$$

The term $n(t, \boldsymbol{x})$ models a random field of noise sources. It is a zero-mean stationary (in time) random process with autocorrelation function

$$\langle n(t_1, \boldsymbol{y}_1) n(t_2, \boldsymbol{y}_2) \rangle = F(t_2 - t_1) \Gamma(\boldsymbol{y}_1, \boldsymbol{y}_2). \tag{1.15}$$

Here $\langle \cdot \rangle$ stands for statistical average with respect to the distribution of the noise sources. For simplicity we will consider that the process n has Gaussian statistics.

The time distribution of the noise sources is characterized by the correlation function $F(t_2 - t_1)$, which is a function of $t_2 - t_1$ only by time stationarity. The function F is normalized so that $F(0) = 1$. The Fourier transform $\hat{F}(\omega)$ of the time correlation function $F(t)$ is a nonnegative, even, real-valued function proportional to the power spectral density of the sources:

$$\hat{F}(\omega) = \int F(t) e^{i\omega t} \mathrm{d}t. \tag{1.16}$$

The spatial distribution of the noise sources is characterized by the autocovariance function $\Gamma(\boldsymbol{y}_1, \boldsymbol{y}_2)$. In most of the book we will assume that the random process n is delta-correlated in space:

$$\Gamma(\boldsymbol{y}_1, \boldsymbol{y}_2) = K(\boldsymbol{y}_1) \delta(\boldsymbol{y}_1 - \boldsymbol{y}_2). \tag{1.17}$$

The function K then characterizes the spatial support of the sources.

1.7 The cross correlation function

The solution of the wave equation has the integral representation (1.8):

$$u(t, \boldsymbol{x}) = \iint n(t - s, \boldsymbol{y}) G(s, \boldsymbol{x}, \boldsymbol{y}) \mathrm{d}s \mathrm{d}\boldsymbol{y},$$

where $G(t, \boldsymbol{x}, \boldsymbol{y})$ is the time-dependent Green's function (1.2).

The empirical cross correlation of the signals recorded at $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ for an integration time T is

$$C_T(\tau, \boldsymbol{x}_1, \boldsymbol{x}_2) = \frac{1}{T} \int_0^T u(t, \boldsymbol{x}_1) u(t + \tau, \boldsymbol{x}_2) \mathrm{d}t. \tag{1.18}$$

It is a statistically stable quantity, in the sense that for a large integration time T, the empirical cross correlation C_T is independent of the realization of the noise sources and it is equal to its expectation. This is stated in the following proposition.

Proposition 1.3. *1. The expectation of the empirical cross correlation C_T (with respect to the distribution of the sources) is independent of T:*

$$\langle C_T(\tau, \boldsymbol{x}_1, \boldsymbol{x}_2)\rangle = C^{(1)}(\tau, \boldsymbol{x}_1, \boldsymbol{x}_2), \tag{1.19}$$

where the statistical cross correlation $C^{(1)}$ *is given by*

$$C^{(1)}(\tau, \boldsymbol{x}_1, \boldsymbol{x}_2) = \frac{1}{2\pi}\iint \mathrm{d}\boldsymbol{y}\mathrm{d}\omega \hat{F}(\omega)K(\boldsymbol{y})\overline{\hat{G}(\omega, \boldsymbol{x}_1, \boldsymbol{y})}\hat{G}(\omega, \boldsymbol{x}_2, \boldsymbol{y})e^{-i\omega\tau}, \tag{1.20}$$

and $\hat{G}(\omega, \boldsymbol{x}, \boldsymbol{y})$ *is the time-harmonic Green's function (i.e. the Fourier transform of* $G(t, \boldsymbol{x}, \boldsymbol{y})$*).*

2. The empirical cross correlation C_T *is a self-averaging quantity:*

$$C_T(\tau, \boldsymbol{x}_1, \boldsymbol{x}_2) \overset{T\to\infty}{\longrightarrow} C^{(1)}(\tau, \boldsymbol{x}_1, \boldsymbol{x}_2), \tag{1.21}$$

in probability with respect to the distribution of the sources.

3. The covariance of the empirical cross correlation C_T *satisfies:*

$$\begin{aligned} &2\pi T\mathrm{Cov}\big(C_T(\tau, \boldsymbol{x}_1, \boldsymbol{x}_2), C_T(\tau + \Delta\tau, \boldsymbol{x}_1, \boldsymbol{x}_2)\big) \\ &\overset{T\to\infty}{\longrightarrow} \int \Big[\int \overline{\hat{G}(\omega, \boldsymbol{x}_1, \boldsymbol{y})}\hat{G}(\omega, \boldsymbol{x}_2, \boldsymbol{y})K(\boldsymbol{y})\mathrm{d}\boldsymbol{y}\Big]^2 \hat{F}(\omega)^2 e^{i\omega(2\tau+\Delta\tau)}\mathrm{d}\omega \\ &+ \int \Big[\int |\hat{G}(\omega, \boldsymbol{x}_1, \boldsymbol{y})|^2 K(\boldsymbol{y})\mathrm{d}\boldsymbol{y}\Big]\Big[\int |\hat{G}(\omega, \boldsymbol{x}_2, \boldsymbol{y})|^2 K(\boldsymbol{y})\mathrm{d}\boldsymbol{y}\Big]\hat{F}(\omega)^2 e^{-i\omega\Delta\tau}\mathrm{d}\omega. \end{aligned} \tag{1.22}$$

1.8 Extended distribution of sources in an inhomogeneous open medium

In this subsection we study the emergence of the Green's function for an extended distribution of sources in an inhomogeneous open medium. Here we give a simple and rigorous proof for an open inhomogeneous medium in the case in which the noise sources are located on the surface of a sphere that encloses both the inhomogeneous region and the sensors, located at $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ (Figure 1.3). The proof is based on the Helmholtz-Kirchhoff identity. It is very simple and convenient, but as we already mentioned, it requires a full aperture illumination. However this condition is not necessary for the cross correlation to be related to the Green's function.

The proof can be summarized as follows: the right-hand side of the Helmholtz-Kirchhoff identity (1.11) is related to the representation (1.20) of the cross correlation function $C^{(1)}$ in the Fourier domain. Therefore, by substituting (1.11) into (1.20) we get the following proposition.

Proposition 1.4. *We assume that*
1) the medium is homogeneous outside the ball $B(\mathbf{0}, D)$ *with center* $\mathbf{0}$ *and radius* D*,*
2) the sources are localized with a uniform density on the sphere $\partial B(\mathbf{0}, L)$ *with center* $\mathbf{0}$ *and radius* L*.*
If $L \gg D$*, then for any* $\boldsymbol{x}_1, \boldsymbol{x}_2 \in B(\mathbf{0}, D)$*, we have*

$$\frac{\partial}{\partial\tau}C^{(1)}(\tau, \boldsymbol{x}_1, \boldsymbol{x}_2) = -\frac{c_0}{2}\left[F * G(\tau, \boldsymbol{x}_1, \boldsymbol{x}_2) - F * G(-\tau, \boldsymbol{x}_1, \boldsymbol{x}_2)\right]. \tag{1.23}$$

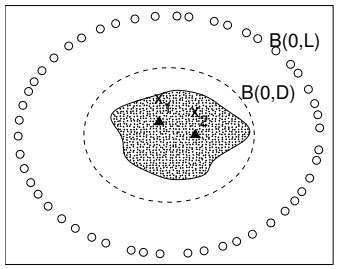


Fig. 1.3. The sources are distributed on the sphere $\partial B(\mathbf{0}, L)$ and the medium is homogeneous outside the ball $B(\mathbf{0}, D)$.

Proof. When the source distribution is uniform over the surface of the ball $B(\mathbf{0}, L)$, the statistical cross correlation (1.20) reads

$$C^{(1)}(\tau, \boldsymbol{x}_1, \boldsymbol{x}_2) = \frac{1}{2\pi} \int \mathrm{d}\omega \hat{F}(\omega) e^{-i\omega\tau} \int_{\partial B(\mathbf{0},L)} \mathrm{d}\sigma(\boldsymbol{y}) \overline{\hat{G}(\omega, \boldsymbol{x}_1, \boldsymbol{y})} \hat{G}(\omega, \boldsymbol{x}_2, \boldsymbol{y}),$$

and its τ-derivative is

$$\begin{aligned} \partial_\tau C^{(1)}(\tau, \boldsymbol{x}_1, \boldsymbol{x}_2) = -\frac{1}{2\pi} \int \mathrm{d}\omega \hat{F}(\omega) e^{-i\omega\tau} \\ \times i\omega \int_{\partial B(\mathbf{0},L)} \mathrm{d}\sigma(\boldsymbol{y}) \hat{F}(\omega) \overline{\hat{G}(\omega, \boldsymbol{x}_1, \boldsymbol{y})} \hat{G}(\omega, \boldsymbol{x}_2, \boldsymbol{y}). \end{aligned}$$

The left-hand side can be simplified by using the Helmholtz-Kirchhoff identity (1.11):

$$\partial_\tau C^{(1)}(\tau, \boldsymbol{x}_1, \boldsymbol{x}_2) = -\frac{c_0}{4\pi} \int \mathrm{d}\omega \hat{F}(\omega) e^{-i\omega\tau} \left[\hat{G}(\omega, \boldsymbol{x}_1, \boldsymbol{x}_2) - \overline{\hat{G}(\omega, \boldsymbol{x}_1, \boldsymbol{x}_2)} \right],$$

which gives the desired result. □

This proposition shows that, when the noise sources surround the region of interest, then the time-lag derivative of the cross correlation of the signals recorded at two observation points is the Green's function between these two points, up to a convolution (in time) with the time covariance function of the noise sources and a symmetrization (which means that we get in fact the causal and the anti-causal Green's functions).

2

Conventional sensor array imaging

In this chapter we give an overview of conventional sensor array imaging for two basic problems: passive source imaging and active reflector imaging.

In the first problem, addressed in Section 2.1, the goal is to image the spatial distribution of sources emitting waves that are recorded by a passive array of receivers. The data set is a vector of N signals recorded by the N receivers. Different imaging functions are introduced. After discussing Least-Squares imaging we introduce the Reverse-Time imaging function and the Kirchhoff Migration imaging function, and we carry out their resolution analysis. The resolution properties are summarized in Section 2.2.3.

In the second problem, addressed in Section 2.3, the goal is to image reflectors buried in the medium from the data collected by an active array of sensors, that can be used both as sources and as receivers. The data set is a matrix of $N \times N$ signals, where the (j,l)-th signal is recorded by the j-th sensor when a short pulse is emitted by the l-th sensor. As in the case of passive source imaging, we discuss Least-Squares imaging, Reverse-Time imaging, and Kirchhoff Migration imaging, whose resolution properties are summarized in Section 2.3.9.

2.1 Passive array imaging of sources

Here we consider the case of a passive array, which means that the sensors are used only as receivers. The goal is to image a source.

2.1.1 Data acquisition

In the configuration described in Figure 2.1, the source $\boldsymbol{y}$ emits a pulse and the sensors $(\boldsymbol{x}_r)_{r=1,\dots,N}$ record the waves. The data set is the vector of signals $\big(u(t,\boldsymbol{x}_r)\big)_{t\in\mathbb{R},r=1,\dots,N}$. The goal of imaging is here to find the source position $\boldsymbol{y}$. More generally, in the case of distributed sources, the goal is to find the source region.

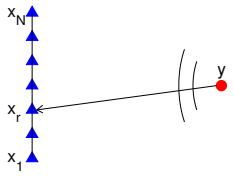


Fig. 2.1. Sensor array imaging of sources.

2.1.2 Imaging function

The goal is to find the spatial source distribution that is supposed to be supported in the region $\Omega \subset \mathbb{R}^3$ (that does not contain the sensor positions $(\boldsymbol{x}_r)_{r=1,\dots,N}$). The data set is $\big(u(t,\boldsymbol{x}_r)\big)_{t\in\mathbb{R},r=1,\dots,N}$. The purpose of imaging is, given the data set, to build an imaging function in the search region Ω:

$$\mathcal{I} : \left| \begin{array}{l} \Omega \to \mathbb{R}^+ \\ \boldsymbol{y}^S \mapsto \mathcal{I}(\boldsymbol{y}^S) \end{array} \right.$$

which plots an image of the source distribution. For instance, in the case in which there is a unique point source, we would like the imaging function to look like a sharp peak centered at the source location. This peak is called point spread function. The width of this peak determines the resolution of the imaging function, which gives the precision with which the source location can be estimated.

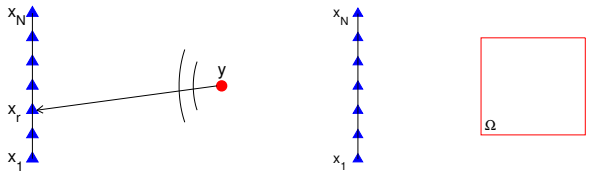


Fig. 2.2. Sensor array imaging of sources: data acquisition (left) and search region for the imaging function (right).

2.1.3 The linear forward operator

The source term is of the form $n(t,\boldsymbol{y}) = \rho_{\text{real}}(\boldsymbol{y})\delta(t)$. The goal is to find the source function ρ_{real}. Here, we assume that the Green's function is known.

The data set is $\hat{\boldsymbol{u}} = \big(\hat{u}(\omega, \boldsymbol{x}_r)\big)_{\omega\in\mathbb{R}, r=1,\dots,N}$ with

$$\hat{u}(\omega, \boldsymbol{x}_r) = \int_\Omega \hat{G}(\omega, \boldsymbol{x}_r, \boldsymbol{y})\rho_{\text{real}}(\boldsymbol{y})\mathrm{d}\boldsymbol{y}.$$

We define the forward operator

$$[\hat{\mathbf{A}}\rho](\omega, \boldsymbol{x}_r) = \int_\Omega \hat{G}(\omega, \boldsymbol{x}_r, \boldsymbol{y})\rho(\boldsymbol{y})\mathrm{d}\boldsymbol{y}. \tag{2.1}$$

$\hat{\mathbf{A}}$ is the linear operator that maps the source function to the array data $\hat{\boldsymbol{u}}$:

$$\hat{\boldsymbol{u}} = \hat{\mathbf{A}}\rho_{\text{real}}.$$

It goes from the space $L^2(\Omega)$ equipped with the standard scalar product

$$(\mu, \nu)_{L^2} = \int_\Omega \overline{\mu(\boldsymbol{y})}\nu(\boldsymbol{y})\mathrm{d}\boldsymbol{y}$$

to $\mathcal{L}^2 := L^2(\mathbb{R}\times\{1,\dots,N\})$ equipped with the scalar product

$$(\hat{\boldsymbol{v}}, \hat{\boldsymbol{w}})_{\mathcal{L}^2} = \int \mathrm{d}\omega \sum_{r=1}^{N} \overline{\hat{v}(\omega, \boldsymbol{x}_r)}\hat{w}(\omega, \boldsymbol{x}_r).$$

2.1.4 The adjoint operator

From the observation of the data we want to estimate ρ_{real}. Note that $\hat{\mathbf{A}}$ is usually not invertible. Therefore we look for the function ρ in $L^2(\Omega)$ that minimizes the misfit between the observed data and the theoretical prediction obtained with ρ: The least squares inverse problem is to minimize $J_{\text{LS}}[\rho]$ where

$$J_{\text{LS}}[\rho] = \frac{1}{2\pi}\int \mathrm{d}\omega \sum_{r=1}^{N} \big|\hat{u}(\omega, \boldsymbol{x}_r) - [\hat{\mathbf{A}}\rho](\omega, \boldsymbol{x}_r)\big|^2. \tag{2.2}$$

As we will see in the next subsection (devoted to least squares inversion), it may be useful to consider the regularized minimization problem with the functional $J_{\text{LS}}[\rho] + \alpha\|\rho\|^2_{\text{REG}}$. Here $\|\rho\|_{\text{REG}}$ is a norm, for instance, the L^2-norm, which smoothes the problem, but which also reduces the resolution. It is also possible to take the L^1-norm, which promotes sparsity in the reconstructed ρ (which is efficient in the case of a few well separated point sources). L^1-minimization techniques (or compressed sensing) have become popular recently in order to reconstruct sparse signals from incomplete data, and in the imaging context it would mean to minimize $\|\rho\|_{L^1}$ such that $\|\hat{\boldsymbol{u}} - \hat{\mathbf{A}}\rho\|_{L^2} \le \epsilon$ for some given positive constant ϵ [13].

The least squares solution to the minimization problem (2.2) solves the normal equation

$$\hat{\mathbf{A}}^*\hat{\mathbf{A}}\rho_{\mathrm{LS}} = \hat{\mathbf{A}}^*\hat{\boldsymbol{u}},$$

where $\hat{\mathbf{A}}^*$ is the adjoint operator

$$\big[\hat{\mathbf{A}}^*\hat{\boldsymbol{v}}\big](\boldsymbol{y}) = \int \mathrm{d}\omega \sum_{r=1}^{N} \overline{\hat{G}(\omega, \boldsymbol{y}, \boldsymbol{x}_r)}\hat{v}(\omega, \boldsymbol{x}_r). \tag{2.3}$$

Indeed, for any $\mu \in L^2(\Omega)$ and $\hat{\boldsymbol{v}} \in \mathcal{L}^2$:

$$\begin{aligned}
\Big(\hat{\mathbf{A}}\mu, \hat{\boldsymbol{v}}\Big)_{\mathcal{L}^2} &= \int \mathrm{d}\omega \sum_{r=1}^{N} \overline{\big[\hat{\mathbf{A}}\mu\big](\omega, \boldsymbol{x}_r)}\hat{v}(\omega, \boldsymbol{x}_r) \\
&= \int \mathrm{d}\omega \sum_{r=1}^{N} \int_{\Omega} \mathrm{d}\boldsymbol{y}\overline{\hat{G}(\omega, \boldsymbol{x}_r, \boldsymbol{y})}\overline{\mu(\boldsymbol{y})}\hat{v}(\omega, \boldsymbol{x}_r) \\
&= \int_{\Omega} \mathrm{d}\boldsymbol{y}\overline{\mu(\boldsymbol{y})} \int \mathrm{d}\omega \sum_{r=1}^{N} \overline{\hat{G}(\omega, \boldsymbol{x}_r, \boldsymbol{y})}\hat{v}(\omega, \boldsymbol{x}_r) \\
&= \int_{\Omega} \mathrm{d}\boldsymbol{y}\overline{\mu(\boldsymbol{y})}\big[\hat{\mathbf{A}}^*\hat{\boldsymbol{v}}\big](\boldsymbol{y}) \quad \text{by Definition (2.3)} \\
&= \Big(\mu, \hat{\mathbf{A}}^*\hat{\boldsymbol{v}}\Big)_{L^2}
\end{aligned}$$

Remember that the complex conjugation in the frequency domain corresponds to the time operation in the time domain. This shows that the adjoint operator corresponds to the backpropagation of the array data to the test point $\boldsymbol{y}$.

The normal operator $\hat{\mathbf{A}}^*\hat{\mathbf{A}}$ is given by

$$\big[\hat{\mathbf{A}}^*\hat{\mathbf{A}}\rho\big](\boldsymbol{y}) = \int_{\Omega} \mathrm{d}\boldsymbol{y}' a(\boldsymbol{y}, \boldsymbol{y}')\rho(\boldsymbol{y}'),$$

with the kernel $a(\boldsymbol{y}, \boldsymbol{y}') = \int \mathrm{d}\omega \sum_{r=1}^{N} \hat{G}(\omega, \boldsymbol{y}, \boldsymbol{x}_r)\overline{\hat{G}(\omega, \boldsymbol{y}', \boldsymbol{x}_r)}$.

2.1.5 Least squares inversion

We consider the problem with a fixed frequency ω to simplify the notations. The data set is the vector $\hat{\boldsymbol{u}} = (\hat{u}(\boldsymbol{x}_r))_{r=1,\dots,N}$. It is related to the unknown function $\rho_{\mathrm{real}} = (\rho_{\mathrm{real}}(\boldsymbol{y}))_{\boldsymbol{y}\in\Omega}$ through the linear relation $\hat{\boldsymbol{u}} = \hat{\mathbf{A}}\rho_{\mathrm{real}}$, where the linear operator $\hat{\mathbf{A}}$ is defined by (2.1). We first put this problem in a simple matrix-vector form. In order to do that, we discretize the problem by introducing a regular grid $(\boldsymbol{y}_j)_{j=1,\dots,M}$ with step δy of the search domain Ω:

$$\big[\hat{\mathbf{A}}\rho\big](\boldsymbol{x}_r) = \sum_{j=1}^{M} \hat{G}(\omega, \boldsymbol{x}_r, \boldsymbol{y}_j)\rho(\boldsymbol{y}_j)(\delta y)^3.$$

The problem is reduced to find the vector $\boldsymbol{\rho} = (\rho(\boldsymbol{y}_j))_{j=1,\dots,M}$ solution of

$$\hat{\boldsymbol{u}} = \mathbf{A}\boldsymbol{\rho}, \quad \text{with the matrix } A_{rj} = \hat{G}(\omega, \boldsymbol{x}_r, \boldsymbol{y}_j)(\delta y)^3.$$

However $\mathbf{A}$ is of size $N \times M$ and even when $M = N$ it is not invertible in general. In order to obtain a well-formulated problem we look for the solution of the least squares inversion problem. It consists in finding the vector $\boldsymbol{\rho}$ that minimizes the error (misfit function):

$$\mathcal{E} = \frac{1}{2}\|\hat{\boldsymbol{u}} - \mathbf{A}\boldsymbol{\rho}\|^2 = \frac{1}{2}\sum_{r=1}^{N} \left|(\hat{\boldsymbol{u}} - \mathbf{A}\boldsymbol{\rho})_r\right|^2.$$

An extremal point of the quadratic misfit function satisfies the constraints:

$$0 = \frac{\partial \mathcal{E}}{\partial \rho_j}, \qquad 0 = \frac{\partial \mathcal{E}}{\partial \overline{\rho_j}},$$

which reads

$$0 = -\sum_{r=1}^{N} \left[\overline{A}_{rj}(\hat{\boldsymbol{u}} - \mathbf{A}\boldsymbol{\rho})_r\right] = -[\mathbf{A}^H(\hat{\boldsymbol{u}} - \mathbf{A}\boldsymbol{\rho})]_j, \quad j = 1, \ldots, N,$$

where the superscript H stands for conjugate transpose. The equation $\mathbf{A}^H(\hat{\boldsymbol{u}} - \mathbf{A}\boldsymbol{\rho}) = \mathbf{0}$ gives the normal equations:

$$\mathbf{A}^H\mathbf{A}\boldsymbol{\rho} = \mathbf{A}^H\hat{\boldsymbol{u}}.$$

$\mathbf{A}^H\mathbf{A}$ is a nonnegative matrix. When it is positive, it can be inverted:

$$\boldsymbol{\rho} = \left(\mathbf{A}^H\mathbf{A}\right)^{-1}\mathbf{A}^H\hat{\boldsymbol{u}}.$$

When $\mathbf{A}^H\mathbf{A}$ is not positive, or when it is positive but ill-conditioned, then it is necessary to regularize the minimization problem. A popular regularization technique is the Tykhonov regularization: the regularized problem to be minimized is

$$\mathcal{E} = \frac{1}{2}\|\hat{\boldsymbol{u}} - \mathbf{A}\boldsymbol{\rho}\|^2 + \frac{1}{2}\alpha\|\boldsymbol{\rho}\|^2,$$

where $\alpha > 0$. The solution to this problem is

$$\boldsymbol{\rho} = \left(\mathbf{A}^H\mathbf{A} + \alpha\mathbf{I}\right)^{-1}\mathbf{A}^H\hat{\boldsymbol{u}}.$$

To be complete, we can add that, when α is small we obtain the Moore-Penrose pseudo-inverse [27], that is known to be unstable and to amplify additive noise. When α is large, the stability is enhanced but the resolution is reduced [10].

2.1.6 The reverse-time imaging function

The least squares imaging function is:

$$\mathcal{I}_{\mathrm{LS}}(\boldsymbol{y}^S) = \left[\left(\hat{\mathbf{A}}^*\hat{\mathbf{A}}\right)^{-1}\hat{\mathbf{A}}^*\hat{\boldsymbol{u}}\right](\boldsymbol{y}^S),$$

where $\left(\hat{\mathbf{A}}^*\hat{\mathbf{A}}\right)^{-1}$ is a (regularized) pseudo-inverse of the normal operator. It turns out that the kernel $a(\boldsymbol{y}, \boldsymbol{y}')$ of the normal operator $\hat{\mathbf{A}}^*\hat{\mathbf{A}}$ is the signal obtained at $\boldsymbol{y}$ during a time-reversal experiment, in which a point source at $\boldsymbol{y}'$ emits a Dirac pulse at time 0 and the time-reversal mirror consists of point receivers/sources at the locations $\boldsymbol{x}_r$, $r = 1, \ldots, N$, as described in Section 1.5. By the time-reversal refocusing property the kernel of the normal operator is often concentrated around $\boldsymbol{y} = \boldsymbol{y}'$, which means that it is close to a diagonal operator. Motivated by this remark we can propose to drop the normalizing factor in the LS function, which gives a new function which is much simpler to evaluate than and reasonably close to the least squares function. This simplification affects, however, resolution [10]. The Reverse-Time imaging function for the search point $\boldsymbol{y}^S$ is defined by:

$$\begin{aligned}\mathcal{I}_{\mathrm{RT}}(\boldsymbol{y}^S) &= \frac{1}{2\pi}\left[\hat{\mathbf{A}}^*\hat{\boldsymbol{u}}\right](\boldsymbol{y}^S) \\ &= \frac{1}{2\pi}\int \mathrm{d}\omega \sum_{r=1}^{N} \overline{\hat{G}(\omega, \boldsymbol{y}^S, \boldsymbol{x}_r)}\hat{u}(\omega, \boldsymbol{x}_r). \end{aligned} \tag{2.4}$$

It is a reverse-time imaging function, since

$$\begin{aligned}\mathcal{I}_{\mathrm{RT}}(\boldsymbol{y}^S) &= \frac{1}{2\pi}\int \mathrm{d}\omega \sum_{r=1}^{N} \overline{\hat{G}(-\omega, \boldsymbol{y}^S, \boldsymbol{x}_r)}\hat{u}(-\omega, \boldsymbol{x}_r) \\ &= \frac{1}{2\pi}\int \mathrm{d}\omega \sum_{r=1}^{N} \hat{G}(\omega, \boldsymbol{y}^S, \boldsymbol{x}_r)\overline{\hat{u}(\omega, \boldsymbol{x}_r)}\end{aligned}$$

can be interpreted as:

$$\mathcal{I}_{\mathrm{RT}}(\boldsymbol{y}^S) = u_{\mathrm{RT}}(0, \boldsymbol{y}^S)$$

where $u_{\mathrm{RT}}(t, \boldsymbol{x})$ is the solution of the wave equation:

$$\frac{1}{c^2(\boldsymbol{x})}\frac{\partial^2 u_{\mathrm{RT}}}{\partial t^2} - \Delta_{\boldsymbol{x}} u_{\mathrm{RT}} = n_{\mathrm{RT}}(t, \boldsymbol{x}),$$

with the source term supported in $(t, \boldsymbol{x}) \in (-\infty, 0) \times \{\boldsymbol{x}_r, r = 1, \ldots, N\}$:

$$n_{\mathrm{RT}}(t, \boldsymbol{x}) = \sum_{r=1}^{N} u(-t, \boldsymbol{x}_r)\delta(\boldsymbol{x} - \boldsymbol{x}_r).$$

This interpretation shows that the computational cost of the imaging function is one call to a solver of the wave equation in the background medium.

2.1.7 Kirchhoff Migration (or travel-time migration)

It is possible to simplify the Reverse-Time imaging function. If we use the geometric optics approximation of the Green's function and neglect the variations of the amplitude term, that is to say, if we make the approximation $\hat{G}(\omega, \boldsymbol{x}, \boldsymbol{y}) \simeq \exp[i\omega\mathcal{T}(\boldsymbol{x}, \boldsymbol{y})]$, where $\mathcal{T}(\boldsymbol{x}, \boldsymbol{y})$ is the travel time from $\boldsymbol{x}$ to $\boldsymbol{y}$, then we get the Kirchhoff Migration imaging function:

$$\begin{aligned}\mathcal{I}_{\mathrm{KM}}(\boldsymbol{y}^S) &= \frac{1}{2\pi}\int \mathrm{d}\omega \sum_{r=1}^{N} \exp[-i\omega\mathcal{T}(\boldsymbol{x}_r, \boldsymbol{y}^S)]\hat{u}(\omega, \boldsymbol{x}_r) \\ &= \sum_{r=1}^{N} u\big(\mathcal{T}(\boldsymbol{x}_r, \boldsymbol{y}^S), \boldsymbol{x}_r\big). \end{aligned} \tag{2.5}$$

Kirchhoff Migration (or travel time migration) has been analyzed in detail [9] and is used extensively in practice. It is a simple way to triangulate the location of a source using sensor array (or distributed sensors).

2.2 Passive array imaging of sources: resolution analysis

In this section we assume that a point source located at $\boldsymbol{y}$ emits a short pulse $f(t)$. The goal of this section is to characterize the resolution properties of the Reverse-Time imaging function, that is to say, the spatial accuracy with which it is possible to localize the source. This can be quantified by the width of the peak of the imaging function at the source position. The Reverse-Time imaging function (2.4) has the form

$$\begin{aligned}\mathcal{I}_{\mathrm{RT}}(\boldsymbol{y}^S) &= \frac{1}{2\pi}\int \hat{\mathcal{I}}_{\mathrm{RT}}(\omega, \boldsymbol{y}^S)\hat{f}(\omega)\mathrm{d}\omega, \\ \hat{\mathcal{I}}_{\mathrm{RT}}(\omega, \boldsymbol{y}^S) &= \sum_{r=1}^{N} \hat{G}(\omega, \boldsymbol{y}^S, \boldsymbol{x}_r)\overline{\hat{G}(\omega, \boldsymbol{x}_r, \boldsymbol{y})}. \end{aligned}$$

We carry out a resolution analysis when the medium is homogeneous. The Green's function is then given by

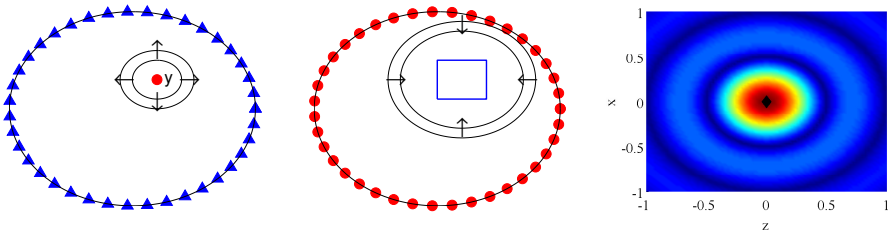
$$\hat{G}(\omega, \boldsymbol{x}, \boldsymbol{y}) = \frac{1}{4\pi|\boldsymbol{x}-\boldsymbol{y}|} e^{i\frac{\omega|\boldsymbol{x}-\boldsymbol{y}|}{c_0}}.$$

2.2.1 Full-aperture array

The resolution analysis in the case of a full-aperture array, that completely surrounds the source, is an application of the analysis of the time-reversal experiment in Section 1.5. It follows from the Helmholtz-Kirchhoff identity (1.11) that gives

$$\hat{\mathcal{I}}_{\mathrm{RT}}(\omega, \boldsymbol{y}^S) \simeq \frac{c_0}{\omega}\mathrm{Im}\big(\hat{G}(\omega, \boldsymbol{y}, \boldsymbol{y}^S)\big) = \frac{1}{4\pi}\mathrm{sinc}\Big(\frac{\omega|\boldsymbol{y}^S - \boldsymbol{y}|}{c_0}\Big),$$

up to a multiplicative constant that depends on the density of sources. We can get the resolution estimate from the first zero (π) of the sinc function. This shows that $\hat{\mathcal{I}}_{\mathrm{RT}}(\omega, \boldsymbol{y}^S)$ is a peak centered at $\boldsymbol{y}$ with width given by $\lambda/2$, with $\lambda = 2\pi c_0/\omega$ the wavelength associated to the frequency ω.



(a) **(b)** **(c)**

Fig. 2.3. A point source at $\boldsymbol{y}$ emits a short pulse that is recorded by a array of receivers that completely surrounds the source (picture a). The RT imaging function backpropagates numerically the time-reversed recorded signals (picture b). The image obtained is a peak centered at the original source location, with a sinc form and a width of $\lambda_0/2$ (picture c, in which x and z are multiples of λ_0 and the modulus of the imaging function is plotted).

If the source has central frequency ω_0 and bandwidth B, with $B \ll \omega_0$,

$$f(t) = \frac{1}{2}e^{-i\omega_0 t} f_0(Bt) + c.c., \tag{2.6}$$

$$\hat{f}(\omega) = \frac{1}{2B}\hat{f}_0\Big(\frac{\omega - \omega_0}{B}\Big) + \frac{1}{2B}\hat{f}_0\Big(\frac{\omega + \omega_0}{B}\Big), \tag{2.7}$$

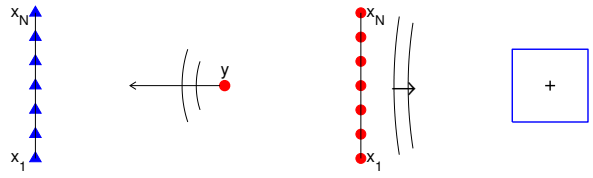
then we have

$$\mathcal{I}_{\mathrm{RT}}(\boldsymbol{y}^S) = \frac{1}{4\pi} f_0(0)\mathrm{sinc}\Big(\frac{\omega_0|\boldsymbol{y}^S - \boldsymbol{y}|}{c_0}\Big).$$

This shows that $\mathcal{I}_{\mathrm{RT}}(\boldsymbol{y}^S)$ is a peak centered at $\boldsymbol{y}$ with width given by $\lambda_0/2$, with $\lambda_0 = 2\pi c_0/\omega_0$ the wavelength associated to the central frequency ω_0 (see Figure 2.3). We have found that the resolution is $\lambda_0/2$, which is a well-known result (diffraction limit).

2.2.2 Partial-aperture array

From now on we consider the case of a finite-aperture array (see Figure 2.4). We denote $\boldsymbol{x} = (\boldsymbol{x}_\perp, z) \in \mathbb{R}^2 \times \mathbb{R}$. We assume that the array lies on the surface $\{z = 0\}$ and that it is dense, with a density function ψ_{r} such that



(a) **(b)**

Fig. 2.4. A point source at $\boldsymbol{y}$ emits a short pulse (picture a). The RT imaging function backpropagates numerically the time-reversed signals to form an image (picture b).

$$\sum_{r=1}^{N} g(\boldsymbol{x}_r) \simeq N \int_{\mathbb{R}^2} \psi_{\mathrm{r}}(\boldsymbol{x}_\perp) g((\boldsymbol{x}_\perp, 0)) \mathrm{d}\boldsymbol{x}_\perp,$$

for any test function g. As a result

$$\hat{\mathcal{I}}_{\mathrm{RT}}(\omega, \boldsymbol{y}^S) = N \int_{\mathbb{R}^2} \psi_{\mathrm{r}}(\boldsymbol{x}_\perp) \hat{G}(\omega, \boldsymbol{y}^S, (\boldsymbol{x}_\perp, 0)) \overline{\hat{G}(\omega, \boldsymbol{y}, (\boldsymbol{x}_\perp, 0))} \mathrm{d}\boldsymbol{x}_\perp. \tag{2.8}$$

We first carry out a high-frequency analysis of the imaging function. We assume that the width of the source pulse $f^\varepsilon(t)$ is small so that

$$f^\varepsilon(t) = f\Big(\frac{t}{\varepsilon}\Big), \qquad \hat{f}^\varepsilon(\omega) = \varepsilon \hat{f}(\varepsilon\omega),$$

where ε is a dimensionless parameter that characterizes the ratio of the pulse width over the typical travel time from the array to the source. Then the imaging function has the form

$$\mathcal{I}_{\mathrm{RT}}(\boldsymbol{y}^S) = \frac{1}{2\pi} \int \hat{\mathcal{I}}_{\mathrm{RT}}(\omega, \boldsymbol{y}^S) \hat{f}(\omega) \mathrm{d}\omega, \tag{2.9}$$

$$\hat{\mathcal{I}}_{\mathrm{RT}}(\omega, \boldsymbol{y}^S) = N \int_{\mathbb{R}^2} a(\boldsymbol{x}_\perp) e^{i \frac{\phi_\omega(\boldsymbol{x}_\perp)}{\varepsilon}} \mathrm{d}\boldsymbol{x}_\perp, \tag{2.10}$$

with

$$a(\boldsymbol{x}_\perp) = \frac{\psi_{\mathrm{r}}(\boldsymbol{x}_\perp)}{16\pi^2 |\boldsymbol{y}^S - (\boldsymbol{x}_\perp, 0)| \, |\boldsymbol{y} - (\boldsymbol{x}_\perp, 0)|},$$
$$\phi_\omega(\boldsymbol{x}_\perp) = \frac{\omega}{c_0} [|\boldsymbol{y}^S - (\boldsymbol{x}_\perp, 0)| - |\boldsymbol{y} - (\boldsymbol{x}_\perp, 0)|].$$

We first give a stationary phase result [38, Chapter VIII].

Lemma 2.1. *Let D is be a compact domain in $\mathbb{R}^2$ and a and ϕ be smooth functions such that there are at most finitely many points in D such that*

$\nabla_\perp \phi(\boldsymbol{x}_\perp) = \mathbf{0}$, *and such that the Hessian of* ϕ *is nondegenerate at these points. Then the value of the integral*

$$I^\varepsilon = \int_D a(\boldsymbol{x}_\perp) e^{i\frac{\phi(\boldsymbol{x}_\perp)}{\varepsilon}} \mathrm{d}\boldsymbol{x}_\perp,$$

is at most of order ε *as* $\varepsilon \to 0$.

As explained in [38], the main contributions to the value of the integral (in the limit $\varepsilon \to 0$) come from critical points, that are stationary points such that $\nabla_\perp \phi(\boldsymbol{x}_\perp) = \mathbf{0}$, and from points on the boundary ∂D such that a level curve of ϕ is tangential to ∂D. This second type of points are absent if a vanishes smoothly at the boundary of D.

Corollary 2.2. *We assume that* $\boldsymbol{y}$ *or* $\boldsymbol{y}^S$ *is not in the plane* $\{z = 0\}$. *In the high-frequency regime* $\varepsilon \to 0$, *we have:*
- if $\boldsymbol{y}^S = \boldsymbol{y}$ *then*

$$\hat{\mathcal{I}}_{\mathrm{RT}}(\omega, \boldsymbol{y}) = \frac{N}{16\pi^2} \int_{\mathbb{R}^2} \frac{\psi_{\mathrm{r}}(\boldsymbol{x}_\perp)}{|(\boldsymbol{x}_\perp, 0) - \boldsymbol{y}|^2} \mathrm{d}\boldsymbol{x}_\perp = O(1), \tag{2.11}$$

- if $\boldsymbol{y}^S \neq \boldsymbol{y}$ *then*

$$\hat{\mathcal{I}}_{\mathrm{RT}}(\omega, \boldsymbol{y}^S) = O(\varepsilon). \tag{2.12}$$

Proof. The result for $\boldsymbol{y}^S = \boldsymbol{y}$ is obvious. Let us assume from now on that $\boldsymbol{y}^S \neq \boldsymbol{y}$ and consider the integral (2.10). The phase function $\phi_\omega(\boldsymbol{x}_\perp)$ is such that

$$\nabla_\perp \phi_\omega(\boldsymbol{x}_\perp) = \frac{\omega}{c_0} \Big[\frac{\boldsymbol{x}_\perp - \boldsymbol{y}_\perp^S}{|\boldsymbol{y}^S - (\boldsymbol{x}_\perp, 0)|} - \frac{\boldsymbol{x}_\perp - \boldsymbol{y}_\perp}{|\boldsymbol{y} - (\boldsymbol{x}_\perp, 0)|} \Big],$$

which is zero if and only if the line going through $\boldsymbol{y}$ and $\boldsymbol{y}^S$ intersects $(\boldsymbol{x}_\perp, 0)$. Therefore, for a given pair of points $\boldsymbol{y} = (\boldsymbol{y}_\perp, L)$ and $\boldsymbol{y}^S = (\boldsymbol{y}_\perp^S, L^S)$, there is at most one stationary point $\boldsymbol{x}_\perp^*$, and in such a case it is nondegenerate as the Hessian of the phase at that point is

$$\mathbf{H}[\phi_\omega](\boldsymbol{x}_\perp^*) = \frac{\omega}{c_0} \frac{\frac{L}{L^S} - 1}{|\boldsymbol{y} - (\boldsymbol{x}_\perp^*, 0)|^3} \begin{pmatrix} L^2 + (x_2^* - y_2)^2 & -(x_1^* - y_1)(x_2^* - y_2) \\ -(x_1^* - y_1)(x_2^* - y_2) & L^2 + (x_1^* - y_1)^2 \end{pmatrix}$$

whose determinant

$$\det \mathbf{H}[\phi_\omega](\boldsymbol{x}_\perp^*) = \frac{\omega^2}{c_0^2} \frac{L^2 (\frac{L}{L^S} - 1)^2}{|\boldsymbol{y} - (\boldsymbol{x}_\perp^*, 0)|^4}$$

is not zero (since $\boldsymbol{y}^S \neq \boldsymbol{y}$ and $\boldsymbol{y}$ and $\boldsymbol{y}^S$ are on the same line issued from $(\boldsymbol{x}_\perp^*, 0)$, we have $L \neq L^S$). The desired result then follows from Lemma 2.1. □

Corollary 2.2 shows that the imaging function has a peak at the original source location. We would like now to characterize the resolution of the imaging function. In the high-frequency regime the result in Corollary 2.2 indicates

that the width of the peak is smaller than one. In fact, the following analysis shows that is is of order ε.

We assume that the third coordinate L of $\boldsymbol{y} = (\boldsymbol{y}_\perp, L)$ is positive. We parameterize the search point $\boldsymbol{y}^S$ as

$$\boldsymbol{y}^S = \boldsymbol{y} + \varepsilon \boldsymbol{z}. \tag{2.13}$$

Then

$$|\boldsymbol{x} - \boldsymbol{y}^S| - |\boldsymbol{x} - \boldsymbol{y}| = -\varepsilon \frac{\boldsymbol{x} - \boldsymbol{y}}{|\boldsymbol{x} - \boldsymbol{y}|} \cdot \boldsymbol{z} + O(\varepsilon^2),$$

and therefore, as $\varepsilon \to 0$,

$$\hat{\mathcal{I}}_{\mathrm{RT}}(\omega, \boldsymbol{y}^S) = \frac{N}{16\pi^2} \int_{\mathbb{R}^2} \frac{\psi_{\mathrm{r}}(\boldsymbol{x}_\perp)}{|(\boldsymbol{x}_\perp, 0) - \boldsymbol{y}|^2} \exp\Big(- i \frac{\omega}{c_0} \frac{\boldsymbol{x} - \boldsymbol{y}}{|\boldsymbol{x} - \boldsymbol{y}|} \cdot \boldsymbol{z}\Big) \mathrm{d}\boldsymbol{x}_\perp. \tag{2.14}$$

We denote $B_2 = \{\boldsymbol{e}_\perp \in \mathbb{R}^2 \text{ s.t. } |\boldsymbol{e}_\perp| < 1\}$ the unit disk in $\mathbb{R}^2$. We introduce the function $\mathcal{X}_{\boldsymbol{y}} : B_2 \to \mathbb{R}^2$ defined by

$$\mathcal{X}_{\boldsymbol{y}}(\boldsymbol{e}_\perp) = \frac{\boldsymbol{e}_\perp}{\sqrt{1 - |\boldsymbol{e}_\perp|^2}} L + \boldsymbol{y}_\perp.$$

The point $(\mathcal{X}_{\boldsymbol{y}}(\boldsymbol{e}_\perp), 0)$ is the intersection of the line starting from $\boldsymbol{y}$ with the direction $(\boldsymbol{e}_\perp, -\sqrt{1 - |\boldsymbol{e}_\perp|^2})$ with the plane $\{z = 0\}$. The inverse of the function $\mathcal{X}_{\boldsymbol{y}}$ is

$$\mathcal{X}_{\boldsymbol{y}}^{-1}(\boldsymbol{x}_\perp) = \frac{\boldsymbol{x}_\perp - \boldsymbol{y}_\perp}{|(\boldsymbol{x}_\perp, 0) - \boldsymbol{y}|},$$

and the determinant of its Jacobian is $|\mathrm{Jac}\, \mathcal{X}_{\boldsymbol{y}}(\boldsymbol{e}_\perp)| = L^2/(1 - |\boldsymbol{e}_\perp|^2)^2$.

A straightforward change of variable $\boldsymbol{x}_\perp \to \mathcal{X}_{\boldsymbol{y}}^{-1}(\boldsymbol{x}_\perp)$ in the expression (2.14) of the reverse-time imaging function gives that, as $\varepsilon \to 0$,

$$\hat{\mathcal{I}}_{\mathrm{RT}}(\omega, \boldsymbol{y}^S) = \frac{N}{16\pi^2} \int_{B_2} \frac{\psi_{\mathrm{r}}(\mathcal{X}_{\boldsymbol{y}}(\boldsymbol{e}_\perp))}{1 - |\boldsymbol{e}_\perp|^2} \exp\Big(i \frac{\omega}{c_0} (-\boldsymbol{e}_\perp, \sqrt{1 - |\boldsymbol{e}_\perp|^2}) \cdot \boldsymbol{z}\Big) \mathrm{d}\boldsymbol{e}_\perp. \tag{2.15}$$

If the density of the array is constant over its support D:

$$\psi_{\mathrm{r}}(\boldsymbol{x}_\perp) = \frac{1}{|D|} \mathbf{1}_D(\boldsymbol{x}_\perp),$$

then we have

$$\hat{\mathcal{I}}_{\mathrm{RT}}(\omega, \boldsymbol{y}^S) = \frac{N}{16\pi^2 |D|} \int_{B_{\boldsymbol{y}}} \frac{1}{1 - |\boldsymbol{e}_\perp|^2} \exp\Big(i \frac{\omega}{c_0} (-\boldsymbol{e}_\perp, \sqrt{1 - |\boldsymbol{e}_\perp|^2}) \cdot \boldsymbol{z}\Big) \mathrm{d}\boldsymbol{e}_\perp, \tag{2.16}$$

where $B_{\boldsymbol{y}} = \mathcal{X}_{\boldsymbol{y}}^{-1}(D)$. The set of unit vectors

$$\mathcal{C}_{\boldsymbol{y}} = \big\{(-\boldsymbol{e}_\perp, \sqrt{1 - |\boldsymbol{e}_\perp|^2}),\, \boldsymbol{e}_\perp \in B_{\boldsymbol{y}}\big\}$$

forms the cone of illumination of the point $\boldsymbol{y}$ (ie the set of directions from a point in the sensor array to the original source location). The function (2.16) gives the local form of the imaging function around the original source location, which is a peak centered at the source location with a width that is of the order of the wavelength. In the following we consider the case of a circular array and of a square array to get some explicit and quantitative formulas that will allow us to distinguish the resolution in the longitudinal direction (z) and in the transverse directions ($\boldsymbol{x}_\perp$).

Circular array

Let us assume that the source location is $\boldsymbol{y} = (0, L)$ and that the support of the sensors is the disk with diameter a:

$$D = \{(\boldsymbol{x}_\perp, 0),\ |\boldsymbol{x}_\perp| \leq a/2\}.$$

Then

$$B_{\boldsymbol{y}} = \Big\{\boldsymbol{e}_\perp \in \mathbb{R}^2,\ |\boldsymbol{e}_\perp| \leq \frac{a}{\sqrt{a^2 + 4L^2}}\Big\}.$$

We take special spherical coordinates to compute (2.16):

$$\boldsymbol{e}_\perp = (\sin\theta\cos\phi, \sin\theta\sin\phi),$$

whose Jacobian determinant is $|\mathrm{Jac}\,\boldsymbol{e}_\perp(\theta,\phi)| = |\sin\theta\cos\theta|$, and we use the identity $\int_0^{2\pi} \exp(ix\sin\phi)\mathrm{d}\phi = 2\pi J_0(x)$ to obtain:

$$\begin{aligned}
&\hat{\mathcal{I}}_{\mathrm{RT}}(\omega, \boldsymbol{y}^S) \\
&= \frac{N}{4\pi^3 a^2}\int_0^{\theta_{\boldsymbol{y}}} \mathrm{d}\theta \int_0^{2\pi} \mathrm{d}\phi \tan\theta \exp\Big(i\frac{\omega}{c_0}\big(-(\cos\phi z_1 + \sin\phi z_2)\sin\theta + \cos\theta z_3\big)\Big) \\
&= \frac{N}{2\pi^2 a^2}\exp\Big(i\frac{\omega}{c_0}z_3\Big)\int_0^{\theta_{\boldsymbol{y}}} \mathrm{d}\theta \tan\theta J_0\Big(\frac{\omega}{c_0}\sin\theta|\boldsymbol{z}_\perp|\Big)\exp\Big(-2i\frac{\omega}{c_0}\sin^2\big(\frac{\theta}{2}\big)z_3\Big),
\end{aligned}$$

with $\theta_{\boldsymbol{y}} = \arctan(a/(2L))$. If we assume that $a \ll L$, then this expression becomes

$$\hat{\mathcal{I}}_{\mathrm{RT}}(\omega, \boldsymbol{y}^S) = \frac{N\theta_{\boldsymbol{y}}^2}{2\pi^2 a^2}\exp\Big(i\frac{\omega}{c_0}z_3\Big)\int_0^1 \mathrm{d}s\, s J_0\Big(\frac{\omega\theta_{\boldsymbol{y}}}{c_0}|\boldsymbol{z}_\perp|s\Big)\exp\Big(-i\frac{\omega\theta_{\boldsymbol{y}}^2}{2c_0}z_3 s^2\Big),$$

or equivalently

$$\hat{\mathcal{I}}_{\mathrm{RT}}(\omega, \boldsymbol{y}^S) = \frac{N}{16\pi^2 L^2}\exp\Big(i\frac{\omega}{c_0}z_3\Big)\Psi\Big(\frac{a\boldsymbol{z}_\perp}{L\lambda}, \frac{a^2 z_3}{L^2\lambda}\Big), \tag{2.17}$$

with the normalized point spread function given by (for $\boldsymbol{\xi}_\perp, \eta \in \mathbb{R}^2 \times \mathbb{R}$):

$$\Psi(\boldsymbol{\xi}_\perp, \eta) = 2\int_0^1 \mathrm{d}s\, s J_0\big(\pi|\boldsymbol{\xi}_\perp|s\big)\exp\Big(-i\pi\frac{\eta}{4}s^2\Big). \tag{2.18}$$

By inspection of formula (2.17-2.18) we observe that:
1) In the transverse directions, the width of the peak is $\lambda L/a$, with $\lambda = 2\pi c_0/\omega$ the wavelength associated to the frequency ω, and the form of the peak is:

$$\begin{aligned}\Psi(\boldsymbol{\xi}_\perp, 0) &= 2\int_0^1 \mathrm{d}s\, s J_0\big(\pi|\boldsymbol{\xi}_\perp|s\big) \\ &= 2\frac{J_1(\pi|\boldsymbol{\xi}_\perp|)}{\pi|\boldsymbol{\xi}_\perp|},\end{aligned}$$

where we have used the derivative identity $\partial_x(xJ_1(x)) = xJ_0(x)$.
2) In the longitudinal direction, the width of the peak is $\lambda L^2/a^2$ and the form of the peak is:

$$\begin{aligned}\Psi(\mathbf{0}, \eta) &= 2\int_0^1 \mathrm{d}s\, s \exp\Big(-i\frac{\pi\eta}{4}s^2\Big) \\ &= 4\frac{1 - e^{-i\frac{\pi\eta}{4}}}{i\pi\eta},\end{aligned}$$

whose square modulus is

$$|\Psi(\mathbf{0}, \eta)|^2 = \mathrm{sinc}^2\Big(\frac{\pi\eta}{8}\Big).$$

If the source has central frequency ω_0 and bandwidth B, with $B \ll \omega_0$, as in (2.6), then we have

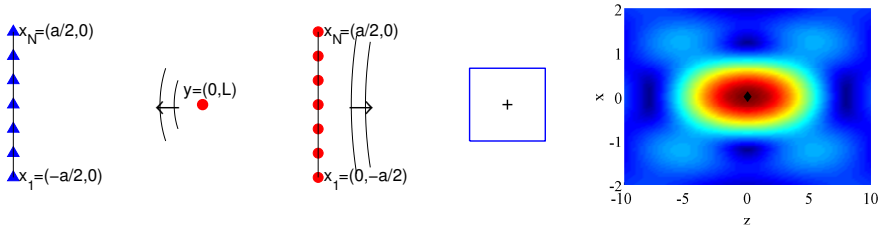
$$\mathcal{I}_{\mathrm{RT}}(\boldsymbol{y}^S) = \frac{N}{32\pi^2 L^2}\exp\Big(i\frac{\omega_0}{c_0}z_3\Big) f_0\Big(\frac{B}{c_0}z_3\Big)\Psi\Big(\frac{a\boldsymbol{z}_\perp}{L\lambda_0}, \frac{a^2 z_3}{L^2\lambda_0}\Big) + c.c. \qquad (2.19)$$

This shows that $\mathcal{I}_{\mathrm{RT}}(\boldsymbol{y}^S)$ is a peak centered at $\boldsymbol{y}$. The transverse width of the peak is $\lambda_0 L/a$, with $\lambda_0 = 2\pi c_0/\omega_0$ the wavelength associated to the central frequency ω_0. This is the so-called Rayleigh resolution formula. The longitudinal width of the peak is $\min\{\lambda_0 L^2/a^2, c_0/B\}$. In other words, if the pulse is narrowband in the sense that $B < \omega_0 a^2/L^2$, then the longitudinal form of the peak centered at the original source location is determined by Fresnel diffraction and it has the form (see Figure 2.5)

$$\mathcal{I}_{\mathrm{RT}}(\boldsymbol{y}^S) = \frac{N f_0(0)}{32\pi^2 L^2}\exp\Big(i\frac{\omega_0}{c_0}z_3\Big)\Psi\Big(\frac{a\boldsymbol{z}_\perp}{L\lambda_0}, \frac{a^2 z_3}{L^2\lambda_0}\Big) + c.c.$$

If the pulse is broadband in the sense that $B > \omega_0 a^2/L^2$, then the longitudinal form of the peak centered at the original source location is determined by the pulse width and it has the form

$$\mathcal{I}_{\mathrm{RT}}(\boldsymbol{y}^S) = \frac{N}{32\pi^2 L^2}\exp\Big(i\frac{\omega_0}{c_0}z_3\Big) f_0\Big(\frac{B}{c_0}z_3\Big)\Psi\Big(\frac{a\boldsymbol{z}_\perp}{L\lambda_0}, 0\Big) + c.c.$$



(a) **(b)** **(c)**

Fig. 2.5. A point source at $\boldsymbol{y} = (0, 0, L)$ emits a short pulse that is recorded by a circular array of receivers in $\{\boldsymbol{x}_\perp \in \mathbb{R}^2, |\boldsymbol{x}_\perp| \leq a/2\} \times \{0\}$ (picture a). The RT imaging function backpropagates numerically the time-reversed recorded signals (picture b). The image obtained is a peak centered at the original source location (picture c, in which x is in multiples of $\lambda_0 L/a$, z in multiples of $\lambda_0 L^2/a^2$, and the modulus of the function is plotted). Here we assume that the bandwidth B of the source pulse is smaller than $\omega_0 a^2/L^2$ so that the form of the peak in the longitudinal direction is not determined by the pulse form but by Fresnel diffraction.

Square array

Let us assume that the source location is $\boldsymbol{y} = (\boldsymbol{0}, L)$ and that the support of the sensors is the square with sidelength a:

$$D = [-a/2, a/2] \times [-a/2, a/2] \times \{0\} = \{(\boldsymbol{x}_\perp, 0),\ |x_1| \leq a/2,\ |x_2| \leq a/2\}.$$

Then

$$B_{\boldsymbol{y}} = \Big\{\boldsymbol{e}_\perp \in \mathbb{R}^2,\ (4L^2 + a^2)e_j^2 \leq a^2(1 - e_{3-j}^2),\ j = 1, 2\Big\},$$

and

$$\hat{\mathcal{I}}_{\rm RT}(\omega, \boldsymbol{y}^S) = \frac{N}{16\pi^2 a^2} \int_{B_{\boldsymbol{y}}} \mathrm{d}e_1 \mathrm{d}e_2 \frac{1}{1 - e_1^2 - e_2^2} \times \exp\Big(i\frac{\omega}{c_0}\big(-(e_1 z_1 + e_2 z_2) + \sqrt{1 - e_1^2 - e_2^2}\, z_3\big)\Big).$$

If we assume that $a \ll L$, then this expression can be simplified:

$$\hat{\mathcal{I}}_{\rm RT}(\omega, \boldsymbol{y}^S) = \frac{N}{16\pi^2 L^2} \exp\Big(i\frac{\omega}{c_0} z_3\Big) \int_{-1/2}^{1/2} \mathrm{d}s_1 \int_{-1/2}^{1/2} \mathrm{d}s_2 \times \exp\Big(-i\frac{\omega a}{c_0 L}(s_1 z_1 + s_2 z_2) - i\frac{\omega a^2}{2c_0 L^2}(s_1^2 + s_2^2) z_3\Big),$$

or equivalently

$$\hat{\mathcal{I}}_{\rm RT}(\omega, \boldsymbol{y}^S) = \frac{N}{16\pi^2 L^2} \exp\Big(i\frac{\omega}{c_0} z_3\Big) \Psi\Big(\frac{a\boldsymbol{z}_\perp}{L\lambda}, \frac{a^2 z_3}{L^2 \lambda}\Big), \tag{2.20}$$

with the normalized point spread function given by

$$\Psi(\boldsymbol{\xi}_\perp,\eta)=\int_{-1/2}^{1/2}\mathrm{d}s_1\int_{-1/2}^{1/2}\mathrm{d}s_2\exp\Big(-2i\pi(s_1\xi_1+s_2\xi_2)-i\pi(s_1^2+s_2^2)\eta\Big). \quad (2.21)$$

By inspection of formula (2.20-2.21) we observe that:
1) In the transverse directions, the width of the peak is $\lambda L/a$ and the form of the peak is:

$$\begin{aligned}\Psi(\boldsymbol{\xi}_\perp,0)&=\int_{-1/2}^{1/2}\mathrm{d}s_1\int_{-1/2}^{1/2}\mathrm{d}s_2\exp\Big(-2i\pi(s_1\xi_1+s_2\xi_2)\Big)\\&=\mathrm{sinc}\big(\pi\xi_1\big)\mathrm{sinc}\big(\pi\xi_2\big).\end{aligned}$$

2) In the longitudinal direction, the width of the peak is $\lambda L^2/a^2$ and the form of the peak is:

$$\begin{aligned}\Psi(\mathbf{0},\eta)&=\int_{-1/2}^{1/2}\mathrm{d}s_1\int_{-1/2}^{1/2}\mathrm{d}s_2\exp\Big(-i\pi(s_1^2+s_2^2)\eta\Big)\\&=\frac{\big(C(\sqrt{\frac{\pi\eta}{4}})-iS(\sqrt{\frac{\pi\eta}{4}})\big)^2}{\frac{\pi\eta}{4}},\end{aligned}$$

where we have used the tabulated Fresnel integrals [1, Section 7.3]:

$$C(x)=\int_0^x\cos(s^2)\mathrm{d}s,\qquad S(x)=\int_0^x\sin(s^2)\mathrm{d}s.$$

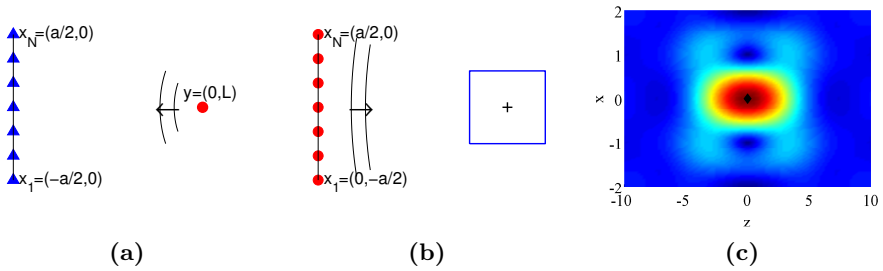


Fig. 2.6. A point source at $\boldsymbol{y}=(0,0,L)$ emits a short pulse that is recorded by a square array of receivers in $[-a/2,a/2]\times[-a/2,a/2]\times\{0\}$ (picture a). The RT imaging function backpropagates numerically the time-reversed recorded signals (picture b). The image obtained is a peak centered at the original source location (picture c, in which x is in multiples of $\lambda_0 L/a$, z in multiples of $\lambda_0 L^2/a^2$, and the modulus of the function is plotted). Here we assume that the bandwidth B of the source pulse is smaller than $\omega_0 a^2/L^2$ so that the form of the peak in the longitudinal direction is not determined by the pulse form but by Fresnel diffraction.

If the source has central frequency ω_0 and bandwidth B, with $B \ll \omega_0$, as in (2.6), then the RT imaging function is given by (2.19) with Ψ given by (2.21). This shows that $\mathcal{I}_{\mathrm{RT}}(\boldsymbol{y}^S)$ is a peak centered at $\boldsymbol{y}$ (see Figure 2.6). The transverse width of the peak is $\lambda_0 L/a$. The longitudinal width of the peak is $\min\{\lambda_0 L^2/a^2, c_0/B\}$.

2.2.3 Summary of resolution analysis for passive source imaging

In this section we have analyzed the resolution properties of the Reverse-Time imaging function. The Kirchhoff Migration imaging function can be analyzed in the same way, as the phase terms are identical, and the results are identical as well.

In the full aperture case, when the passive array completely surrounds the source, the latter can be localized with a precision of the order of the diffraction limit, that is, of the order of the central wavelength λ.

In the partial aperture case, when the passive array has a diameter a and the distance from the array to the source is L, the latter can be localized in the transverse direction with a precision of the order of the Rayleigh resolution formula, that is, of the order of $\lambda L/a$, where λ is the central wavelength. It can be localized in the longitudinal direction with a precision of the order of $\min\{\lambda L^2/a^2, c_0/B\}$, where B is the bandwidth of the source.

The resolution formulas that we have described in this section are well known and presented in many books. For instance, the analysis of the transverse resolution formulas (for arrays of various forms) can be found in [11, Section 8.5] and the analysis of the longitudinal ones in [11, Section 8.8].

2.3 Active array imaging of reflectors

In this section we assume that the sensor array is active, which means that the sensors can be used as sources and/or as receivers. The goal is to find a reflector buried into the medium.

2.3.1 Data acquisition

As described in Figure 2.7, in which the sensor array consists of N sensors, the data acquisition is carried out in N steps. For each $s = 1, \ldots, N$, the source $\boldsymbol{x}_s$ emits a pulse and the sensor at $\boldsymbol{x}_r$ (for $r = 1, \ldots, N$) record the wave that we denote by $u(t, \boldsymbol{x}_r, \boldsymbol{x}_s)$. The data set is the time-dependent response matrix $(u(t, \boldsymbol{x}_r, \boldsymbol{x}_s))_{t\in\mathbb{R}, r,s=1,\ldots,N}$, also called impulse response matrix. The goal of imaging is here to find the reflector position $\boldsymbol{y}$ (more generally, find the reflectivity function of the medium).

The data set is $(u(t, \boldsymbol{x}_r, \boldsymbol{x}_s))_{t\in\mathbb{R}, r,s=1,\ldots,N}$. Given the data set, we wish to build an imaging function in the search region $\Omega \subset \mathbb{R}^3$:

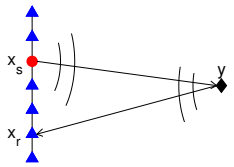


Fig. 2.7. Sensor array imaging of reflectors.

$$\mathcal{I} : \left| \begin{array}{l} \Omega \to \mathbb{R}^+ \\ \boldsymbol{y}^S \mapsto \mathcal{I}(\boldsymbol{y}^S) \end{array} \right.$$

which plots an image of the reflectivity function in the search region.

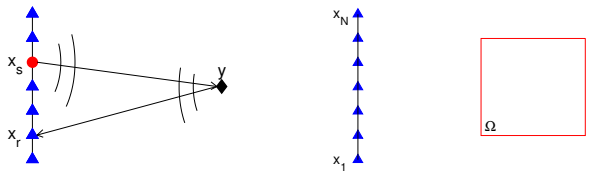


Fig. 2.8. Sensor array imaging of reflectors: data acquisition (left) and search region for the imaging function (right).

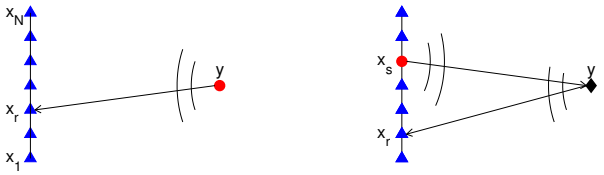
2.3.2 Source and reflector array imaging: comparison

Source and reflector sensor array imaging are two different processes. In source array imaging the goal is to image an unknown source given the vector of waves recorded by the array. In reflector array imaging the goal is to image an unknown reflectivity (i.e. the contrast in the speed of propagation) given the matrix of waves emitted and recorded by the array:

2.3.3 Modeling

The goal is to find the propagation speed $(c_{\text{real}}(\boldsymbol{y}))_{\boldsymbol{y}\in\Omega}$ which is assumed to be homogeneous outside the domain Ω. The Fourier transforms of the recorded signals are

$$\hat{u}(\omega, \boldsymbol{x}_r, \boldsymbol{x}_s) = \hat{G}(\omega, \boldsymbol{x}_r, \boldsymbol{x}_s; c_{\text{real}})\hat{f}(\omega),$$



Source array of sensors
The sensors $(\boldsymbol{x}_r)_{r=1,\dots,N}$ record
$\boldsymbol{y}$ is a source
Data: $\big(u(t,\boldsymbol{x}_r)\big)_{t\in\mathbb{R},r=1,\dots,N}$

Reflector array of sensors/sources
The sensors $(\boldsymbol{x}_r)_{r=1,\dots,N}$ record
$\boldsymbol{y}$ is a reflector, $\boldsymbol{x}_s$ is a source
Data: $\big(u(t,\boldsymbol{x}_r,\boldsymbol{x}_s)\big)_{t\in\mathbb{R},r,s=1,\dots,N}$

Fig. 2.9. Comparison between sensor array imaging of sources (left) and of reflectors (right).

where $\hat{f}(\omega)$ is the Fourier transform of the source pulse used by the array and $\hat{G}(\omega,\boldsymbol{x},\boldsymbol{y};c)$ is the Green's function that solves the Helmholtz equation (1.4) with the propagation speed $c(\boldsymbol{x})$ and with the Sommerfeld radiation condition (1.5). Here we explicitly write the dependence of the Green's function on the velocity $c(\boldsymbol{x})$. Note that this dependence is nonlinear.

2.3.4 Nonlinear inversion

The data set is the matrix of signals:

$$\hat{\mathbf{u}} = \big(\hat{u}(\omega,\boldsymbol{x}_r,\boldsymbol{x}_s)\big)_{\omega\in\mathbb{R},r,s=1,\dots,N} = \big(\hat{G}(\omega,\boldsymbol{x}_r,\boldsymbol{x}_s;c_{\text{real}})\hat{f}(\omega)\big)_{\omega\in\mathbb{R},r,s=1,\dots,N}.$$

In order to estimate the propagation speed $c_{\text{real}}(\boldsymbol{x})$, we try so solve the array least squares inverse problem, that is to say, to find the function c that minimizes:

$$J[c] + \alpha\|c\|^2_{\text{REG}},$$

where $J[c]$ is the misfit function

$$J[c] = \frac{1}{2\pi}\int \mathrm{d}\omega \sum_{r,s=1}^{N} \big|\hat{u}(\omega,\boldsymbol{x}_r,\boldsymbol{x}_s) - \hat{G}(\omega,\boldsymbol{x}_r,\boldsymbol{x}_s;c)\hat{f}(\omega)\big|^2,$$

α is a strength of regularization parameter, and $\|c\|_{\text{REG}}$ is a norm used to regularize the least squares problem. Note that, in contrast to the passive imaging case, this is a nonlinear problem for the unknown propagation speed $c(\boldsymbol{x})$ and the misfit function is not quadratic in the unknown. It is, in general, very difficult to solve this problem. Iterative techniques are possible, but the solution is often very unstable and depends strongly on the initial guess. In the seismic imaging framework, the resolution of the least squares inverse problem is called Full Waveform Inversion.

2.3.5 Linearization of the forward problem

We consider the following model for the speed of propagation:

$$\frac{1}{c^2(\boldsymbol{x})} = \frac{1}{c_0^2}\big(n_0^2(\boldsymbol{x}) + \rho(\boldsymbol{x})\big), \tag{2.22}$$

where c_0 is a reference speed (known), $n_0(\boldsymbol{x})$ is a smooth background index of refraction (known, typically constant), and $\rho(\boldsymbol{x})$ is the target reflectivity (unknown but assumed to be small). The goal of reflector imaging is to reconstruct ρ given the data set.

The Green's function satisfies:

$$\Delta_{\boldsymbol{x}}\hat{G} + \frac{\omega^2}{c_0^2}\big(n_0^2(\boldsymbol{x}) + \rho(\boldsymbol{x})\big)\hat{G} = -\delta(\boldsymbol{x}-\boldsymbol{y}).$$

The background Green's function is the solution of:

$$\Delta_{\boldsymbol{x}}\hat{G}_0 + \frac{\omega^2}{c_0^2}n_0^2(\boldsymbol{x})\hat{G}_0 = -\delta(\boldsymbol{x}-\boldsymbol{y}),$$

with the Sommerfeld radiation condition. In the Born approximation we have the approximation:

$$\hat{G}(\omega,\boldsymbol{x},\boldsymbol{y}) = \hat{G}_0(\omega,\boldsymbol{x},\boldsymbol{y}) + \frac{\omega^2}{c_0^2}\int \hat{G}_0(\omega,\boldsymbol{x},\boldsymbol{z})\rho(\boldsymbol{z})\hat{G}_0(\omega,\boldsymbol{z},\boldsymbol{y})\mathrm{d}\boldsymbol{z}, \tag{2.23}$$

in which the first term of the right-hand side corresponds to the direct waves (i.e. those which have not interacted with the reflector) and the second term corresponds to the single-scattered waves emitted from the source point $\boldsymbol{y}$, scattered at $\boldsymbol{z}$ (in the support of ρ) and received at $\boldsymbol{x}$. The Born approximation is valid when the reflectivity of the reflector is small enough.

Proof. Let us consider the full Green's function and the background Green's function solution of

$$\Delta_{\boldsymbol{z}}\hat{G}(\omega,\boldsymbol{z},\boldsymbol{x}) + \frac{\omega^2}{c_0^2}n_0^2(\boldsymbol{z})\hat{G}(\omega,\boldsymbol{z},\boldsymbol{x}) = -\frac{\omega^2}{c_0^2}\rho(\boldsymbol{z})\hat{G}(\omega,\boldsymbol{z},\boldsymbol{x}) - \delta(\boldsymbol{z}-\boldsymbol{x}),$$
$$\Delta_{\boldsymbol{z}}\hat{G}_0(\omega,\boldsymbol{z},\boldsymbol{y}) + \frac{\omega^2}{c_0^2}n_0^2(\boldsymbol{z})\hat{G}_0(\omega,\boldsymbol{z},\boldsymbol{y}) = -\delta(\boldsymbol{z}-\boldsymbol{y}).$$

We multiply the first equation by $\hat{G}_0(\omega,\boldsymbol{x},\boldsymbol{y})$ and subtract the second equation multiplied by $\hat{G}(\omega,\boldsymbol{x},\boldsymbol{z})$:

$$
\begin{aligned}
&\nabla_{\boldsymbol{z}} \cdot \left[\hat{G}_0(\omega,\boldsymbol{z},\boldsymbol{y})\nabla_{\boldsymbol{z}}\hat{G}(\omega,\boldsymbol{z},\boldsymbol{x}) - \hat{G}(\omega,\boldsymbol{z},\boldsymbol{y})\nabla_{\boldsymbol{z}}\hat{G}_0(\omega,\boldsymbol{z},\boldsymbol{x})\right] \\
&= -\frac{\omega^2}{c_0^2}\rho(\boldsymbol{z})\hat{G}(\omega,\boldsymbol{z},\boldsymbol{x})\hat{G}_0(\omega,\boldsymbol{z},\boldsymbol{y}) - \hat{G}_0(\omega,\boldsymbol{z},\boldsymbol{y})\delta(\boldsymbol{z}-\boldsymbol{x}) + \hat{G}(\omega,\boldsymbol{z},\boldsymbol{x})\delta(\boldsymbol{z}-\boldsymbol{y}) \\
&= -\frac{\omega^2}{c_0^2}\rho(\boldsymbol{z})\hat{G}(\omega,\boldsymbol{z},\boldsymbol{x})\hat{G}_0(\omega,\boldsymbol{z},\boldsymbol{y}) - \hat{G}_0(\omega,\boldsymbol{x},\boldsymbol{y})\delta(\boldsymbol{z}-\boldsymbol{x}) + \hat{G}(\omega,\boldsymbol{y},\boldsymbol{x})\delta(\boldsymbol{z}-\boldsymbol{y}) \\
&\overset{reciprocity}{=} -\frac{\omega^2}{c_0^2}\rho(\boldsymbol{z})\hat{G}(\omega,\boldsymbol{x},\boldsymbol{z})\hat{G}_0(\omega,\boldsymbol{z},\boldsymbol{y}) - \hat{G}_0(\omega,\boldsymbol{x},\boldsymbol{y})\delta(\boldsymbol{z}-\boldsymbol{x}) \\
&\qquad\qquad +\hat{G}(\omega,\boldsymbol{x},\boldsymbol{y})\delta(\boldsymbol{z}-\boldsymbol{y}).
\end{aligned}
$$

We integrate over $B(\mathbf{0}, L)$ (with L large enough so that it encloses the support of ρ):

$$
0 = -\frac{\omega^2}{c_0^2}\int \hat{G}(\omega,\boldsymbol{x},\boldsymbol{z})\rho(\boldsymbol{z})\hat{G}_0(\omega,\boldsymbol{z},\boldsymbol{y})\mathrm{d}\boldsymbol{z} - \hat{G}_0(\omega,\boldsymbol{x},\boldsymbol{y}) + \hat{G}(\omega,\boldsymbol{x},\boldsymbol{y}).
$$

We therefore obtain the Lippmann-Schwinger equation, which is exact:

$$
\hat{G}(\omega,\boldsymbol{x},\boldsymbol{y}) = \hat{G}_0(\omega,\boldsymbol{x},\boldsymbol{y}) + \frac{\omega^2}{c_0^2}\int \hat{G}(\omega,\boldsymbol{x},\boldsymbol{z})\rho(\boldsymbol{z})\hat{G}_0(\omega,\boldsymbol{z},\boldsymbol{y})\mathrm{d}\boldsymbol{z}. \tag{2.24}
$$

This equation is used as a basis for expanding the Green's functions when the reflectivity ρ is small. If the full Green's function $\hat{G}$ in the right-hand side is replaced by the background Green's function, then we obtain:

$$
\hat{G}(\omega,\boldsymbol{x},\boldsymbol{y}) \simeq \hat{G}_0(\omega,\boldsymbol{x},\boldsymbol{y}) + \frac{\omega^2}{c_0^2}\int \hat{G}_0(\omega,\boldsymbol{x},\boldsymbol{z})\rho(\boldsymbol{z})\hat{G}_0(\omega,\boldsymbol{z},\boldsymbol{y})\mathrm{d}\boldsymbol{z},
$$

which is the (first-order) Born approximation. □

The data set is modeled by $\hat{\mathbf{u}} = (\hat{u}(\omega,\boldsymbol{x}_r,\boldsymbol{x}_s))_{\omega\in\mathbb{R},r,s=1,\dots,N}$ with

$$
\hat{u}(\omega,\boldsymbol{x}_r,\boldsymbol{x}_s) = \frac{\omega^2}{c_0^2}\int \hat{G}_0(\omega,\boldsymbol{x}_r,\boldsymbol{z})\rho(\boldsymbol{z})\hat{G}_0(\omega,\boldsymbol{y},\boldsymbol{x}_s)\mathrm{d}\boldsymbol{z}. \tag{2.25}
$$

Note that we have removed $\hat{G}_0(\omega,\boldsymbol{x}_r,\boldsymbol{x}_s)\hat{f}(\omega)$ from the original data set, and then rescaled by $\hat{f}(\omega)$ for ω in the bandwidth of the source (this procedure is known as equalization), which is possible since they are known quantities. We define the forward operator

$$
[\hat{\mathbf{A}}\rho](\omega,\boldsymbol{x}_r,\boldsymbol{x}_s) = \int_\Omega \hat{G}_0(\omega,\boldsymbol{x}_r,\boldsymbol{z})\rho(\boldsymbol{z})\hat{G}_0(\omega,\boldsymbol{z},\boldsymbol{x}_s)\mathrm{d}\boldsymbol{z}. \tag{2.26}
$$

It is the linear operator that maps the reflectivity function to the array data. It goes from the space $L^2(\Omega)$ equipped with the standard scalar product

$$
(\mu,\nu)_{L^2} = \int_\Omega \overline{\mu(\boldsymbol{y})}\nu(\boldsymbol{y})\mathrm{d}\boldsymbol{y}
$$

to $\mathcal{L}^2 := L^2(\mathbb{R} \times \{1, \ldots, N\}^2)$ equipped with the scalar product

$$(\hat{\mathbf{v}}, \hat{\mathbf{w}})_{\mathcal{L}^2} = \int \mathrm{d}\omega \sum_{r,s=1}^{N} \overline{\hat{v}(\omega, \boldsymbol{x}_r, \boldsymbol{x}_s)} \hat{w}(\omega, \boldsymbol{x}_r, \boldsymbol{x}_s).$$

2.3.6 Linearized inversion

The least squares linearized inverse problem consists in minimizing $J_{\mathrm{LS}}[\rho]$ where

$$J_{\mathrm{LS}}[\rho] = \frac{1}{2\pi} \int \mathrm{d}\omega \sum_{r,s=1}^{N} \left| \hat{u}(\omega, \boldsymbol{x}_r, \boldsymbol{x}_s) - [\hat{\mathbf{A}}\rho](\omega, \boldsymbol{x}_r, \boldsymbol{x}_s) \right|^2.$$

The solution of the least squares linearized inverse problem solves the normal equations:

$$\left(\hat{\mathbf{A}}^* \hat{\mathbf{A}}\right) \rho_{\mathrm{LS}} = \hat{\mathbf{A}}^* \hat{\mathbf{u}}.$$

Here the adjoint operator is

$$[\hat{\mathbf{A}}^* \hat{\mathbf{v}}](\boldsymbol{y}) = \sum_{r,s=1}^{N} \int \mathrm{d}\omega \overline{\hat{G}_0(\omega, \boldsymbol{y}, \boldsymbol{x}_r) \hat{G}_0(\omega, \boldsymbol{x}_s, \boldsymbol{y})} \hat{v}(\omega, \boldsymbol{x}_r, \boldsymbol{x}_s).$$

Remember that the complex conjugation in the frequency domain corresponds to the time-reversal operation in the time domain. This shows that the adjoint operator corresponds to the backpropagation of the array data both from the receiver point $\boldsymbol{x}_r$ and from the source point $\boldsymbol{x}_s$ to the test point $\boldsymbol{y}$. The normal operator is

$$\left[\hat{\mathbf{A}}^* \hat{\mathbf{A}} \rho\right](\boldsymbol{y}) = \int_{\Omega} \mathrm{d}\boldsymbol{y}' a(\boldsymbol{y}, \boldsymbol{y}') \rho(\boldsymbol{y}'),$$

with the kernel

$$a(\boldsymbol{y}, \boldsymbol{y}') = \int \mathrm{d}\omega \sum_{r,s=1}^{N} \hat{G}_0(\omega, \boldsymbol{y}, \boldsymbol{x}_r) \hat{G}_0(\omega, \boldsymbol{x}_s, \boldsymbol{y}) \overline{\hat{G}_0(\omega, \boldsymbol{y}', \boldsymbol{x}_r) \hat{G}_0(\omega, \boldsymbol{x}_s, \boldsymbol{y}')}.$$

As a result, the least squares imaging function is:

$$\mathcal{I}_{\mathrm{LS}}(\boldsymbol{y}^S) = \left[\left(\hat{\mathbf{A}}^* \hat{\mathbf{A}}\right)^{-1} \hat{\mathbf{A}}^* \hat{\boldsymbol{u}} \right](\boldsymbol{y}^S),$$

where $\left(\hat{\mathbf{A}}^* \hat{\mathbf{A}}\right)^{-1}$ is a (regularized) pseudo-inverse of the normal operator.

2.3.7 The reverse-time imaging function

Motivated again by the fact that $\hat{\mathbf{A}}^* \hat{\mathbf{A}}$ is often close to a diagonal operator, we suggest to drop this term to get a simplified imaging function. The Reverse-Time imaging function for the search point $\boldsymbol{y}^S$ is defined by:

$$\begin{aligned}\mathcal{I}_{\mathrm{RT}}(\boldsymbol{y}^S) &= \frac{1}{2\pi}\big[\hat{\mathbf{A}}^*\hat{\mathbf{u}}\big](\boldsymbol{y}^S)\\ &= \frac{1}{2\pi}\int \mathrm{d}\omega \sum_{r,s=1}^{N} \overline{\hat{G}_0(\omega,\boldsymbol{y}^S,\boldsymbol{x}_r)\hat{G}_0(\omega,\boldsymbol{x}_s,\boldsymbol{y}^S)}\hat{u}(\omega,\boldsymbol{x}_r,\boldsymbol{x}_s)\\ &= \frac{1}{2\pi}\int \mathrm{d}\omega \sum_{r,s=1}^{N} \hat{G}_0(\omega,\boldsymbol{y}^S,\boldsymbol{x}_r)\hat{G}_0(\omega,\boldsymbol{x}_s,\boldsymbol{y}^S)\overline{\hat{u}(\omega,\boldsymbol{x}_r,\boldsymbol{x}_s)}. \quad (2.27)\end{aligned}$$

The reverse-time imaging function can be evaluated as follows:
- Solve for $s = 1, \ldots, N$ the wave equation in the background medium

$$\frac{n_0^2(\boldsymbol{x})}{c_0^2}\frac{\partial^2 u_{\mathrm{RT}}^{(s)}}{\partial t^2} - \Delta_{\boldsymbol{x}} u_{\mathrm{RT}}^{(s)} = n_{\mathrm{RT}}^{(s)}(t,\boldsymbol{x}),$$

with the source term supported in $(t,\boldsymbol{x}) \in (-\infty,0)\times\{\boldsymbol{x}_r, r=1,\ldots,N\}$:

$$n_{\mathrm{RT}}^{(s)}(t,\boldsymbol{x}) = \sum_{r=1}^{N}\delta(\boldsymbol{x}-\boldsymbol{x}_r)u(-t,\boldsymbol{x}_r;\boldsymbol{x}_s).$$

- Solve for $s = 1, \ldots, N$ the wave equation in the background medium

$$\frac{n_0^2(\boldsymbol{x})}{c_0^2}\frac{\partial^2 w^{(s)}}{\partial t^2} - \Delta_{\boldsymbol{x}} w^{(s)} = n^{(s)}(t,\boldsymbol{x}),$$

with the source term:

$$n^{(s)}(t,\boldsymbol{x}) = \delta(\boldsymbol{x}-\boldsymbol{x}_s)\delta(t).$$

- Evaluate for any search point $\boldsymbol{y}^S$ the correlation

$$\widetilde{\mathcal{I}}_{\mathrm{RT}}(\boldsymbol{y}^S) = \sum_{s=1}^{N}\int_0^\infty w^{(s)}(t,\boldsymbol{y}^S)u_{\mathrm{RT}}^{(s)}(-t,\boldsymbol{y}^S)\mathrm{d}t.$$

This algorithm gives indeed the desired result because the wave solutions can be expressed as

$$\begin{aligned}\hat{u}_{\mathrm{RT}}^{(s)}(\omega,\boldsymbol{x}) &= \sum_{r=1}^{N}\hat{G}_0(\omega,\boldsymbol{x},\boldsymbol{x}_r)\overline{\hat{u}(\omega,\boldsymbol{x}_r,\boldsymbol{x}_s)},\\ \hat{w}^{(s)}(\omega,\boldsymbol{x}) &= \hat{G}_0(\omega,\boldsymbol{x},\boldsymbol{x}_s),\end{aligned}$$

and therefore the correlation of the wave solutions is equal to the reverse-time imaging function:

$$\begin{aligned}\widetilde{\mathcal{I}}_{\mathrm{RT}}(\boldsymbol{y}^S) &= \sum_{s=1}^{N}\int_{-\infty}^{\infty} w^{(s)}(t,\boldsymbol{y}^S)u_{\mathrm{RT}}^{(s)}(-t,\boldsymbol{y}^S)\mathrm{d}t \\ &= \frac{1}{2\pi}\sum_{s=1}^{N}\int_{-\infty}^{\infty} \hat{w}^{(s)}(\omega,\boldsymbol{y}^S)\hat{u}_{\mathrm{RT}}^{(s)}(\omega,\boldsymbol{y}^S)\mathrm{d}\omega \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{d}\omega\sum_{r,s=1}^{N}\hat{G}_0(\omega,\boldsymbol{y}^S,\boldsymbol{x}_r)\hat{G}_0(\omega,\boldsymbol{x}_s,\boldsymbol{y}^S)\overline{\hat{u}(\omega,\boldsymbol{x}_r,\boldsymbol{x}_s)} \\ &= \mathcal{I}_{\mathrm{RT}}(\boldsymbol{y}^S).\end{aligned}$$

This interpretation shows that the computational cost of the reverse-time imaging function is $2N$ calls to a solver of the wave equation in the background medium with $2N$ sets of different sources (here we neglect the cost of the evaluation of the correlations).

2.3.8 Kirchhoff migration (or travel-time migration)

The Kirchhoff migration (or travel time migration) is obtained as a simplification of the reverse-time imaging function in which we take $\hat{G}_0(\omega,\boldsymbol{x},\boldsymbol{y}) \simeq \exp[i\omega\mathcal{T}(\boldsymbol{x},\boldsymbol{y})]$, where $\mathcal{T}(\boldsymbol{x},\boldsymbol{y})$ is the travel time from $\boldsymbol{x}$ to $\boldsymbol{y}$. Therefore the Kirchhoff migration imaging function has the form:

$$\begin{aligned}\mathcal{I}_{\mathrm{KM}}(\boldsymbol{y}^S) &= \frac{1}{2\pi}\int\mathrm{d}\omega\sum_{r,s=1}^{N}\exp[-i\omega(\mathcal{T}(\boldsymbol{x}_r,\boldsymbol{y}^S)+\mathcal{T}(\boldsymbol{x}_s,\boldsymbol{y}^S))]\hat{u}(\omega,\boldsymbol{x}_r,\boldsymbol{x}_s) \\ &= \sum_{r,s=1}^{N} u\big(\mathcal{T}(\boldsymbol{x}_r,\boldsymbol{y}^S)+\mathcal{T}(\boldsymbol{x}_s,\boldsymbol{y}^S),\boldsymbol{x}_r,\boldsymbol{x}_s\big). \qquad (2.28)\end{aligned}$$

Kirchhoff Migration (or travel time migration) has been analyzed in detail [9]. It is a simple way to triangulate the location of a reflector using an active sensor array.

2.3.9 Summary of resolution analysis for active reflector imaging

The resolution analysis of the Reverse-Time imaging function and the Kirchhoff Migration imaging function goes along the same way as for passive source imaging.

In the full aperture case, when the active array completely surrounds the reflector, the latter can be localized with a precision of the order of the diffraction limit, that is, of the order of the central wavelength λ.

In the partial aperture case, when the active array has a diameter a and the distance from the array to the reflector is L, the latter can be localized in the transverse direction with a precision of the order of the Rayleigh resolution formula, that is, of the order of $\lambda L/a$, where λ is the central wavelength. It

can be localized in the longitudinal direction with a precision of the order of $\min\{\lambda L^2/a^2, c_0/B\}$, where B is the bandwidth of the pulse emitted by the array sensors.

Remark : In seismology the spatial resolution of migrated images has been studied extensively, including situations with slowly varying background, in [7, 16]. These studies confirm that the transverse and longitudinal resolution is a function of the array aperture, the distance from the array to the reflector, and the spectral bandwidth of the illuminating wave field.

2.4 A remark about time-reversal experiments

Originally time reversal was proposed not for imaging, but for energy focusing. The idea was to focus ultrasound energy on kidney stones to destroy them [21]. In a time-reversal experiment the Time-Reversal Mirror (TRM) is used first as a receiver array, then as a source array.

The source $\boldsymbol{y}$ emits a pulse
The TRM records the signals

The TRM emits the time-reversed signals
A sensor array probes the region
around the original source location

Fig. 2.10. The two steps of a time-reversal experiment.

It is striking that time reversal looks like reverse-time migration for passive array imaging. In both cases the recorded data are backpropagated from the array. There is, however, a difference: backpropagation is performed physically in a Time-Reversal experiment and numerically in Reverse-Time migration:
- In Reverse-Time migration, backpropagation is carried out numerically, in a fictitious medium (corresponding to the known background medium).
- In Time Reversal, backpropagation is carried out physically, in the real medium.

There is no difference when the medium is perfectly known (and the numerical code is perfect), but the difference will have dramatic consequences when the medium becomes complex and partially known.

3

Time-harmonic Reverse-Time imaging with measurement noise

3.1 The data set

Let us consider the propagation of scalar waves in a three-dimensional medium. In the presence of a localized reflector, the speed of propagation can be modeled by

$$\frac{1}{c^2(\boldsymbol{x})} = \frac{1}{c_o^2}\big(1 + V_{\text{ref}}(\boldsymbol{x})\big). \tag{3.1}$$

Here
- the constant c_o is the known background speed,
- the local variation $V_{\text{ref}}(\boldsymbol{x})$ of the speed of propagation induced by the reflector at $\boldsymbol{z}_{\text{ref}}$ is of the form

$$V_{\text{ref}}(\boldsymbol{x}) = \sigma_{\text{ref}} \mathbf{1}_{\Omega_{\text{ref}}}(\boldsymbol{x} - \boldsymbol{z}_{\text{ref}}), \tag{3.2}$$

where σ_{ref} is the reflectivity of the reflector, $\boldsymbol{z}_{\text{ref}}$ is its center, and Ω_{ref} is a compactly supported domain with volume l_{ref}^3 that models its spatial support.

Suppose that we have co-localized time-harmonic transmitter and receiver arrays $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$ of n elements, used to detect the reflector. In the presence of a reflector, the field received by the jth receiving element $\boldsymbol{x}_j$ when the transmitter at $\boldsymbol{x}_l$ emits a unitary time-harmonic wave is $\hat{u}(\omega, \boldsymbol{x}_j, \boldsymbol{x}_l)$, where $\hat{u}(\omega, \boldsymbol{x}, \boldsymbol{y})$ the solution to the Helmholtz equation

$$\Delta_{\boldsymbol{x}} \hat{u} + \frac{\omega^2}{c^2(\boldsymbol{x})} \hat{u} = -\delta(\boldsymbol{x} - \boldsymbol{y}), \tag{3.3}$$

with the Sommerfeld radiation condition imposed on $\hat{u}$.

The data set collected by the array describes the transmit-receive process performed at this array (see Figure 3.1). If we remove the incident field then it can be defined as

$$\big\{\hat{v}(\omega, \boldsymbol{x}_j, \boldsymbol{x}_l),\ j, l = 1, \ldots, n\big\}, \tag{3.4}$$

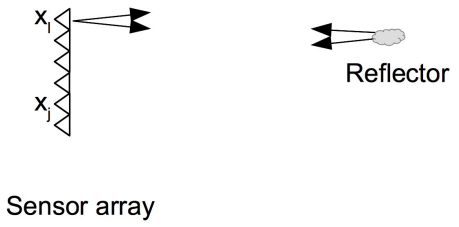


Fig. 3.1. Reverse-Time imaging set-up. $\boldsymbol{x}_l$ emits a time-harmonic wave that propagates through the medium, is backscattered by the reflector, and is recorded by the sensor at $\boldsymbol{x}_j$.

with

$$\hat{v}(\omega, \boldsymbol{x}_j, \boldsymbol{x}_l) = \hat{u}(\omega, \boldsymbol{x}_j, \boldsymbol{x}_l) - \hat{G}_0(\omega, \boldsymbol{x}_j, \boldsymbol{x}_l) + \hat{W}_{j,l}. \tag{3.5}$$

Here $\hat{W}_{j,l}$ represents the additive measurement noise. The incident field is given in terms of the homogeneous Green's function $\hat{G}_0(\omega, \boldsymbol{x}, \boldsymbol{y})$ defined by

$$\hat{G}_0(\omega, \boldsymbol{x}, \boldsymbol{y}) = \frac{1}{4\pi|\boldsymbol{x}-\boldsymbol{y}|} \exp\Big(i\frac{\omega}{c_o}|\boldsymbol{x}-\boldsymbol{y}|\Big), \tag{3.6}$$

that is the solution to

$$\Delta_{\boldsymbol{x}}\hat{G}_0(\omega, \boldsymbol{x}, \boldsymbol{y}) + \frac{\omega^2}{c_o^2}\hat{G}_0(\omega, \boldsymbol{x}, \boldsymbol{y}) = -\delta(\boldsymbol{x}-\boldsymbol{y}), \tag{3.7}$$

with the Sommerfeld radiation condition.

3.2 The forward problem

A reflector is embedded at $\boldsymbol{z}_{\text{ref}}$ and is modeled by the local variation $V_{\text{ref}}(\boldsymbol{x})$ of the propagation speed as in (4.1). The full Green's function $\hat{G}$, that is to say, the Green's function of the medium in the presence of the reflector at $\boldsymbol{z}_{\text{ref}}$, is solution of

$$\Delta_{\boldsymbol{x}}\hat{G}\big(\omega, \boldsymbol{x}, \boldsymbol{y}\big) + \frac{\omega^2}{c^2(\boldsymbol{x})}\hat{G}\big(\omega, \boldsymbol{x}, \boldsymbol{y}\big) = -\delta(\boldsymbol{x}-\boldsymbol{y}). \tag{3.8}$$

The Lippmann-Schwinger integral equation for the full Green's function $\hat{G}$ is

$$\begin{aligned}\hat{G}(\omega, \boldsymbol{x}, \boldsymbol{y}) &= \hat{G}_0(\omega, \boldsymbol{x}, \boldsymbol{y}) \\ &\quad + \frac{\omega^2}{c_o^2}\int \hat{G}_0(\omega, \boldsymbol{x}, \boldsymbol{z})V_{\text{ref}}(\boldsymbol{z})\hat{G}(\omega, \boldsymbol{z}, \boldsymbol{y})\mathrm{d}\boldsymbol{z}.\end{aligned} \tag{3.9}$$

Using the Born approximation, we get

$$\hat{G}(\omega, \boldsymbol{x}, \boldsymbol{y}) = \hat{G}_0(\omega, \boldsymbol{x}, \boldsymbol{y}) + \frac{\omega^2}{c_o^2} \int \hat{G}_0(\omega, \boldsymbol{x}, \boldsymbol{z}) V_{\mathrm{ref}}(\boldsymbol{z}) \hat{G}_0(\omega, \boldsymbol{z}, \boldsymbol{y}) \mathrm{d}\boldsymbol{z}. \tag{3.10}$$

Therefore the full Green's function can be written as the sum

$$\hat{G}(\omega, \boldsymbol{x}, \boldsymbol{y}) = \hat{G}_0(\omega, \boldsymbol{x}, \boldsymbol{y}) + \hat{G}_{\mathrm{ref}}(\omega, \boldsymbol{x}, \boldsymbol{y}). \tag{3.11}$$

The term $\hat{G}_{\mathrm{ref}}$ is the term in the data set that corresponds to the reflector:

$$\hat{G}_{\mathrm{ref}}(\omega, \boldsymbol{x}, \boldsymbol{y}) = \frac{\omega^2}{c_o^2} \int \hat{G}_0(\omega, \boldsymbol{x}, \boldsymbol{z}) V_{\mathrm{ref}}(\boldsymbol{z}) \hat{G}_0(\omega, \boldsymbol{z}, \boldsymbol{y}) \mathrm{d}\boldsymbol{z}.$$

The approximation (3.11) is formally valid if the correction $\hat{G}_{\mathrm{ref}}$ is small compared to $\hat{G}_0$, i.e., in the regime in which $\sigma_{\mathrm{ref}} \ll 1$, with an error that is formally of order $O(\sigma_{\mathrm{ref}}^2)$. We also assume that the diameter l_{ref} of the scattering region Ω_{ref} is small compared to the typical wavelength. We can then model the reflector by a point reflector (the point interaction approximation)

$$V_{\mathrm{ref}}(\boldsymbol{x}) \approx \sigma_{\mathrm{ref}} l_{\mathrm{ref}}^3 \delta(\boldsymbol{x} - \boldsymbol{z}_{\mathrm{ref}}), \tag{3.12}$$

and we can write the correction in the form

$$\hat{G}_{\mathrm{ref}}(\omega, \boldsymbol{x}, \boldsymbol{y}) = \frac{\omega^2}{c_o^2} \sigma_{\mathrm{ref}} l_{\mathrm{ref}}^3 \hat{G}_0(\omega, \boldsymbol{x}, \boldsymbol{z}_{\mathrm{ref}}) \hat{G}_0(\omega, \boldsymbol{z}_{\mathrm{ref}}, \boldsymbol{y}). \tag{3.13}$$

The data set is therefore of the form

$$\hat{v}(\omega, \boldsymbol{x}_j, \boldsymbol{x}_l) = \hat{v}_{\mathrm{ref}}(\omega, \boldsymbol{x}_j, \boldsymbol{x}_l) + \hat{W}_{jl}, \tag{3.14}$$

with $\hat{v}_{\mathrm{ref}}$ the ideal data set in the absence of measurement noise

$$\begin{aligned} \hat{v}_{\mathrm{ref}}(\omega, \boldsymbol{x}_j, \boldsymbol{x}_l) &= \hat{G}_{\mathrm{ref}}\big(\omega, \boldsymbol{x}_j, \boldsymbol{x}_l\big) \\ &= \frac{\omega^2}{c_o^2} \sigma_{\mathrm{ref}} l_{\mathrm{ref}}^3 \hat{G}_0\big(\omega, \boldsymbol{x}_j, \boldsymbol{z}_{\mathrm{ref}}\big) \hat{G}_0\big(\omega, \boldsymbol{z}_{\mathrm{ref}}, \boldsymbol{x}_l\big), \end{aligned} \tag{3.15}$$

and $\hat{W}_{jl}$ is the measurement noise. We assume that the random variables $\hat{W}_{jl}$ are independent and identically distributed, with a circular complex Gaussian distribution with variance σ_{mes}^2, that is to say, $\mathrm{Re}(\hat{W}_{jl})$ and $\mathrm{Im}(\hat{W}_{jl})$ are independent and identically distributed real-valued Gaussian random variables with mean zero and variance $\sigma_{\mathrm{mes}}^2/2$.

3.3 The imaging function

In the presence of a point reflector at $\boldsymbol{z}_{\mathrm{ref}}$ and in the presence of additive noise the data set is of the form (3.14). We study the effect of the measurement noise on the time-harmonic reverse-time imaging function defined by

$$\mathcal{I}_{\mathrm{RT}}(\boldsymbol{x}) = \frac{1}{n^2} \sum_{j,l=1}^{n} \hat{G}_0(\omega, \boldsymbol{x}, \boldsymbol{x}_j) \hat{G}_0(\omega, \boldsymbol{x}, \boldsymbol{x}_l) \overline{\hat{v}(\omega, \boldsymbol{x}_j, \boldsymbol{x}_l)}. \tag{3.16}$$

3.3.1 The imaging function without measurement noise

In the absence of noise $\sigma_{\text{mes}} = 0$ the imaging function is given by

$$\mathcal{I}_{\text{RT}}(\boldsymbol{x}) = \mathcal{I}_0(\boldsymbol{x}), \qquad \mathcal{I}_0(\boldsymbol{x}) = \sigma_{\text{ref}} l_{\text{ref}}^3 \frac{\omega^2}{c_o^2} \mathcal{H}(\boldsymbol{x}, \boldsymbol{z}_{\text{ref}}), \tag{3.17}$$

where

$$\mathcal{H}(\boldsymbol{x}, \boldsymbol{y}) = \Big(\frac{1}{n} \sum_{j=1}^{n} \overline{\hat{G}_0(\omega, \boldsymbol{x}, \boldsymbol{x}_j)} \hat{G}_0(\omega, \boldsymbol{y}, \boldsymbol{x}_j) \Big)^2. \tag{3.18}$$

The function $\boldsymbol{x} \to \mathcal{H}(\boldsymbol{x}, \boldsymbol{z}_{\text{ref}})$ is the point spread function that describes the spatial profile of the peak obtained at the reflector location $\boldsymbol{z}_{\text{ref}}$ in the imaging function when the reflector is point-like.

Full-aperture array. If the sensor array is dense (i.e. the inter-sensor distance is smaller than half a wavelength) and completely surrounds the region of interest (for instance, it covers the surface of the ball with center at $\mathbf{0}$ and radius L) then Helmholtz-Kirchhoff theorem states that $\mathcal{H}(\boldsymbol{x}, \boldsymbol{y})$ is proportional to the square of the imaginary part of the Green's function $\hat{G}_0(\omega, \boldsymbol{x}, \boldsymbol{y})$. We find

$$\mathcal{H}(\boldsymbol{x}, \boldsymbol{y}) = C_L \widetilde{\mathcal{H}}(\boldsymbol{x} - \boldsymbol{y}), \text{ where } \widetilde{\mathcal{H}}(\boldsymbol{x}) = \operatorname{sinc}^2\Big(\frac{\omega |\boldsymbol{x}|}{c_o} \Big), \tag{3.19}$$

and $C_L = 1/(4\pi L)^4$. Therefore we have

$$\mathcal{I}_0(\boldsymbol{x}) = \sigma_0 \widetilde{\mathcal{H}}(\boldsymbol{x} - \boldsymbol{z}_{\text{ref}}), \tag{3.20}$$

where

$$\sigma_0 = \frac{\sigma_{\text{ref}} l_{\text{ref}}^3 \omega^2}{(4\pi L)^4 c_o^2}, \qquad \widetilde{\mathcal{H}}(\boldsymbol{x}) = \operatorname{sinc}^2\Big(\frac{2\pi |\boldsymbol{x}|}{\lambda} \Big), \tag{3.21}$$

and $\lambda = 2\pi c_o/\omega$ is the wavelength. This shows that the width of the point spread function $\widetilde{\mathcal{H}}(\boldsymbol{x})$ is of the order of $\lambda/2$, which is the Abbe diffraction limit [11].

3.3.2 The imaging function with measurement noise

In the presence of measurement noise the imaging function is a complex Gaussian random field. Its mean is the unperturbed imaging function $\mathcal{I}_0$ defined by (3.17), its relation function is zero:

$$\mathbb{E}\big[\big(\mathcal{I}_{\text{RT}}(\boldsymbol{x}) - \mathcal{I}_0(\boldsymbol{x}) \big) \big(\mathcal{I}_{\text{RT}}(\boldsymbol{x}') - \mathcal{I}_0(\boldsymbol{x}') \big) \big] = 0,$$

and its covariance functions is:

$$\begin{aligned}\mathrm{Cov}\big(\mathcal{I}_{\mathrm{RT}}(\boldsymbol{x}),\mathcal{I}_{\mathrm{RT}}(\boldsymbol{x}')\big) &= \mathbb{E}\big[\big(\mathcal{I}_{\mathrm{RT}}(\boldsymbol{x})-\mathcal{I}_0(\boldsymbol{x})\big)\overline{\big(\mathcal{I}_{\mathrm{RT}}(\boldsymbol{x}')-\mathcal{I}_0(\boldsymbol{x}')\big)}\big],\\ &= \frac{\sigma_{\mathrm{mes}}^2}{n^2}\Big(\frac{1}{n}\sum_{j=1}^{n}\hat{G}_0(\omega,\boldsymbol{x},\boldsymbol{x}_j)\overline{\hat{G}_0(\omega,\boldsymbol{x}',\boldsymbol{x}_j)}\Big)^2\\ &= \frac{\sigma_{\mathrm{mes}}^2}{n^2}\mathcal{H}(\boldsymbol{x},\boldsymbol{x}'). \end{aligned} \tag{3.22}$$

If we assume that $\mathcal{H}(\boldsymbol{x},\boldsymbol{y}) = C_L\widetilde{\mathcal{H}}(\boldsymbol{x}-\boldsymbol{y})$, with $\widetilde{\mathcal{H}}(\boldsymbol{x})$ a smooth peaked function centered at $\mathbf{0}$, as in the case of a full-aperture array discussed above, then we can see that the messurement noise creates a speckle noise $\mathcal{I}_{\mathrm{RT}}-\mathcal{I}_0$ in the image, which is a stationary Gaussian random field with mean zero, variance

$$\sigma_{\mathrm{noise}}^2 = \frac{\sigma_{\mathrm{mes}}^2}{n^2}C_L\widetilde{\mathcal{H}}(\mathbf{0}), \tag{3.23}$$

and covariance function:

$$\mathrm{Cov}\big(\mathcal{I}_{\mathrm{RT}}(\boldsymbol{x}),\mathcal{I}_{\mathrm{RT}}(\boldsymbol{x}')\big) = \sigma_{\mathrm{noise}}^2\mathcal{H}_{\mathrm{noise}}(\boldsymbol{x}-\boldsymbol{x}'), \tag{3.24}$$

with

$$\mathcal{H}_{\mathrm{noise}}(\boldsymbol{x}) = \frac{\widetilde{\mathcal{H}}(\boldsymbol{x})}{\widetilde{\mathcal{H}}(\mathbf{0})}. \tag{3.25}$$

This random field is a speckle pattern whose hotspot profiles are close to the function $\mathcal{H}_{\mathrm{noise}}(\boldsymbol{x})$, which is (proportional to) the point spread function of the imaging function. Here the hotspot profile refers to the local shape of the field around a local maximum (see Appendix 6.5).

The hotspot volume is defined as

$$V_c = \frac{4\pi^2}{(\det\boldsymbol{\Lambda})^{1/2}}, \qquad \boldsymbol{\Lambda} = \Big(-\partial^2_{x_jx_l}\mathcal{H}_{\mathrm{noise}}(\boldsymbol{x})\,|_{\boldsymbol{x}=\mathbf{0}}\Big)_{j,l=1,\dots,3}. \tag{3.26}$$

The maximum of the random field $\mathcal{I}_{\mathrm{RT}}(\boldsymbol{x})-\mathcal{I}_0(\boldsymbol{x})$ over a domain Ω whose volume is much larger than the hotspot volume is a random quantity described in Appendix 6.5, which is equal to a deterministic value to leading order in $|\Omega|/V_c$:

$$\max_{\boldsymbol{x}\in\Omega}\big\{\mathcal{I}_{\mathrm{RT}}(\boldsymbol{x})-\mathcal{I}_0(\boldsymbol{x})\big\} = \sigma_{\mathrm{noise}}\sqrt{2\log\Big(\frac{|\Omega|}{V_c}\Big)}. \tag{3.27}$$

Full-aperture array. In the case in which the array completely surrounds the region of interest, $\mathcal{I}_{\mathrm{RT}}(\boldsymbol{x})$ is a Gaussian random field with mean $\mathcal{I}_0(\boldsymbol{x})$ given by (4.35), variance

$$\sigma_{\mathrm{noise}}^2 = \frac{\sigma_{\mathrm{mes}}^2}{n^2}\frac{1}{(4\pi L)^4}, \tag{3.28}$$

and covariance function:

$$\mathrm{Cov}\big(\mathcal{I}_{\mathrm{RT}}(\boldsymbol{x}),\mathcal{I}_{\mathrm{RT}}(\boldsymbol{x}')\big)=\sigma^2_{\mathrm{noise}}\mathcal{H}_{\mathrm{noise}}(\boldsymbol{x}-\boldsymbol{x}'),\qquad \mathcal{H}_{\mathrm{noise}}(\boldsymbol{x})=\mathrm{sinc}^2\Big(\frac{2\pi|\boldsymbol{x}|}{\lambda}\Big). \tag{3.29}$$

The speckle pattern is made of hotspots with typical radius λ and typical amplitude σ_{noise}. The typical shape of the hotspot is given by the function $\mathcal{H}_{\mathrm{noise}}(\boldsymbol{x})$, that has a slow power law decay as $1/|\boldsymbol{x}|^2$. The signal to noise ratio in the image is

$$\mathrm{SNR}=\frac{\mathcal{I}_0(\boldsymbol{z}_{\mathrm{ref}})}{\sigma_{\mathrm{noise}}}=\frac{\sigma_0}{\sigma_{\mathrm{noise}}}=\frac{n\sigma_{\mathrm{ref}}l^3_{\mathrm{ref}}\omega^2}{\sigma_{\mathrm{mes}}c_o^2(4\pi L)^2}=\frac{n\sigma_{\mathrm{ref}}l^3_{\mathrm{ref}}}{4\sigma_{\mathrm{mes}}\lambda^2L^2}. \tag{3.30}$$

Note that:
- The SNR increases with the number of sensors, and this is because the additive noise is independent from one sensor to the other one.
- The SNR decays with the square of the wavelength, because the scattering efficiency (and therefore the reflected signal amplitude) at small wavelengths is inversely proportional to the square of the wavelength.
- The SNR decays with the square of the distance from the array to the reflector, because the reflected signal amplitude is inversely proportional to the distance from the array to the reflector (in a three-dimensional homogeneous medium).

In the case of the full-aperture array, the matrix $\boldsymbol{\Lambda}$ is proportional to the identity and the hotspot volume is proportional to the wavelength to the power three

$$\boldsymbol{\Lambda}=\frac{8\pi^2}{3\lambda^2}\mathbf{I},\qquad V_c=\frac{3^{3/2}}{2^{5/2}\pi}\lambda^3. \tag{3.31}$$

This result shows that the SNR (3.30) should be considered with cautious. If σ_{noise} is of the same order as σ_0, then the speckle pattern may have a local maximum whose peak amplitude is much larger than σ_{noise} and that could be misinterpreted as a reflector. In [5] a detailed statistical analysis was carried out in order to build a detection test with maximal power of detection for a given fals alarm rate.

3.3.3 Localization error

The localization of the reflector consists in looking after the maximum of the imaging functional (the statistical approach proposed in [5] shows that the location of the maximum of the Reverse-Time imaging function is the Maximum Likelihood Estimator of the location of the reflector). In the presence of a reflector at $\boldsymbol{z}_{\mathrm{ref}}$, we have seen that the imaging functional has the form

$$\mathcal{I}_{\mathrm{RT}}(\boldsymbol{x})=\mathcal{I}_0(\boldsymbol{x})+\mathcal{I}_1(\boldsymbol{x}),$$

where $\mathcal{I}_0$ is the unperturbed imaging function given by (3.17) and $\mathcal{I}_1$ is a complex Gaussian random field with mean zero, variance $\sigma^2_{\mathrm{noise}}$, and covariance function $\sigma^2_{\mathrm{noise}}\mathcal{H}_{\mathrm{noise}}(\boldsymbol{x}-\boldsymbol{x}')$.

We consider the case in which $\mathcal{H}_{\text{noise}}$ is real-valued, which is the case for a full-aperture array, and we denote by ℓ the width of the point spread function, which is of the order of λ for a full-aperture array. We assume $\sigma_{\text{noise}} \ll \sigma_0$, so that a Taylor series expansion around $\boldsymbol{z}_{\text{ref}}$, for $|\boldsymbol{x}-\boldsymbol{z}_{\text{ref}}| \lesssim \ell/\text{SNR} = \ell\sigma_{\text{noise}}/\sigma_0$, gives:

$$\mathcal{I}_{\text{RT}}(\boldsymbol{x}) = \sigma_0\Big(1 - \frac{1}{2}(\boldsymbol{x}-\boldsymbol{z}_{\text{ref}})^T\boldsymbol{\Lambda}(\boldsymbol{x}-\boldsymbol{z}_{\text{ref}}) + \frac{1}{\sigma_0}\mathcal{I}_1(\boldsymbol{z}_{\text{ref}}) + \frac{1}{\sigma_0}\nabla\mathcal{I}_1(\boldsymbol{z}_{\text{ref}})^T(\boldsymbol{x}-\boldsymbol{z}_{\text{ref}}) + O\big(\frac{\sigma_{\text{noise}}^3}{\sigma_0^3}\big)\Big). \tag{3.32}$$

It was shown in [5] that the maximum likelihood estimator of the location of the reflector is the location of the maximum of the Reverse-Time imaging function:

$$\hat{\boldsymbol{z}} = \operatorname*{argmax}_{\boldsymbol{x}} \big|\mathcal{I}_{\text{RT}}(\boldsymbol{x})\big|^2.$$

It has the form:

$$\hat{\boldsymbol{z}} = \boldsymbol{z}_{\text{ref}} + \frac{1}{\sigma_0}\text{Re}\big(\boldsymbol{\Lambda}^{-1}\nabla\mathcal{I}_1(\boldsymbol{z}_{\text{ref}})\big) + O\big(\ell\frac{\sigma_{\text{noise}}^2}{\sigma_0^2}\big).$$

To leading order (in $\sigma_{\text{noise}}/\sigma_0$) the estimator $\hat{\boldsymbol{z}}$ is unbiased, i.e. its mean is the true location $\boldsymbol{z}_{\text{ref}}$. Moreover, using the fact that $\mathbb{E}\big[\text{Re}\nabla\mathcal{I}_1(\boldsymbol{z}_{\text{ref}})\text{Re}\nabla\mathcal{I}_1(\boldsymbol{z}_{\text{ref}})^T\big] = \sigma_{\text{noise}}^2\boldsymbol{\Lambda}/2$, the covariance matrix of the estimator $\hat{\boldsymbol{z}}$ is

$$\mathbb{E}\big[(\hat{\boldsymbol{z}}-\boldsymbol{z}_{\text{ref}})(\hat{\boldsymbol{z}}-\boldsymbol{z}_{\text{ref}})^T\big] = \frac{\sigma_{\text{noise}}^2}{2\sigma_0^2}\boldsymbol{\Lambda}^{-1} = \frac{1}{2\,\text{SNR}^2}\boldsymbol{\Lambda}^{-1}, \tag{3.33}$$

which is order ℓ^2/SNR^2. This means that the relative error (relative to the radius ℓ of the point spread function) in the localization of the reflector is of the order of $1/\text{SNR} = \sigma_{\text{noise}}/\sigma_0$. Note also that, as a byproduct of this analysis, we find that the perturbed value of the maximum of the peak is of the form

$$\big|\mathcal{I}_{\text{RT}}(\hat{\boldsymbol{z}})\big|^2 \simeq \sigma_0^2\Big(1 + \frac{2}{\sigma_0}\text{Re}\big(\mathcal{I}_1(\boldsymbol{z}_{\text{ref}})\big) + O\big(\frac{\sigma_{\text{noise}}^2}{\sigma_0^2}\big)\Big), \tag{3.34}$$

where $\text{Re}\big(\mathcal{I}_1(\boldsymbol{z}_{\text{ref}})\big)$ follows a Gaussian distribution with mean 0 and variance $\sigma_{\text{noise}}^2/2$.

Full-aperture array. If the sensor array is dense and surrounds the region of interest, then the localization errors are independent along the three directions and their variances are

$$\mathbb{E}\big[(\hat{z}_j - z_{\text{ref},j})^2\big] = \frac{\sigma_{\text{noise}}^2}{\sigma_0^2}\frac{3\lambda^2}{16\pi^2} = \frac{1}{\text{SNR}^2}\frac{3\lambda^2}{16\pi^2}, \quad j = 1,\dots,3.$$

This formula shows that the resolution is proportional to the wavelength and inversely proportional to the signal-to-noise ratio in the image (3.30). This means that signal-to-noise ratio and resolution are related to each other. When the noise level is vanishing and there is a point reflector, or when there are well separated point reflectors, i.e. when the object to be imaged is sparse, then it is possible to localize the reflectors with infinite resolution.

4

Time-Harmonic Reverse-Time imaging with clutter noise

4.1 The data set

Let us consider the propagation of scalar waves in a three-dimensional medium. In the presence of a localized reflector and small random fluctuations of the medium, the speed of propagation can be modeled by

$$\frac{1}{c^2(\boldsymbol{x})} = \frac{1}{c_o^2}\big(1 + V_{\text{clu}}(\boldsymbol{x}) + V_{\text{ref}}(\boldsymbol{x})\big). \tag{4.1}$$

Here
- the constant c_o is the known background speed,
- the random process $V_{\text{clu}}(\boldsymbol{x})$ represents the cluttered medium,
- the local variation $V_{\text{ref}}(\boldsymbol{x})$ of the speed of propagation induced by the reflector at $\boldsymbol{z}_{\text{ref}}$ is of the form

$$V_{\text{ref}}(\boldsymbol{x}) = \sigma_{\text{ref}} \mathbf{1}_{\Omega_{\text{ref}}}(\boldsymbol{x} - \boldsymbol{z}_{\text{ref}}), \tag{4.2}$$

where σ_{ref} is the reflectivity of the reflector, $\boldsymbol{z}_{\text{ref}}$ is its center, and Ω_{ref} is a compactly supported domain with volume l_{ref}^3 that models its spatial support.

Suppose that we have co-localized transmitter and receiver arrays $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$ of n elements, used to detect the reflector. In the presence of a reflector and small random fluctuations of the medium, the field received by the jth receiving element $\boldsymbol{x}_j$ when the pulse $f(t)$ is emitted from $\boldsymbol{x}_l$ is $u(t, \boldsymbol{x}_j, \boldsymbol{x}_l)$, where $(t, \boldsymbol{x}) \to u(t, \boldsymbol{x}, \boldsymbol{y})$ is the solution to the scalar wave equation

$$\frac{1}{c^2(\boldsymbol{x})}\frac{\partial^2 u}{\partial t^2} - \Delta_{\boldsymbol{x}} u = f(t)\delta(\boldsymbol{x} - \boldsymbol{y}), \tag{4.3}$$

or, in the Fourier domain, $(\omega, \boldsymbol{x}) \to \hat{u}(\omega, \boldsymbol{x}, \boldsymbol{y})$ is the solution to the Helmholtz equation

$$\Delta_{\boldsymbol{x}} \hat{u} + \frac{\omega^2}{c^2(\boldsymbol{x})}\hat{u} = -\hat{f}(\omega)\delta(\boldsymbol{x} - \boldsymbol{y}), \tag{4.4}$$

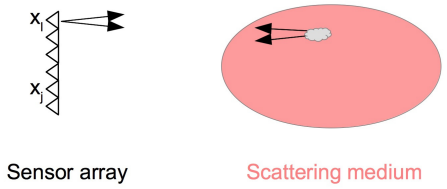


Fig. 4.1. Reverse-Time imaging set-up. $\boldsymbol{x}_l$ emits a short pulse that propagates through the scattering medium, is backscattered by the reflector, and is recorded by the sensor at $\boldsymbol{x}_j$.

with the Sommerfeld radiation condition imposed on $\hat{u}$.

The data set collected by the array describes the transmit-receive process performed at this array (see Figure 4.1). If we remove the incident field then it can be defined as

$$\big\{\hat{v}(\omega, \boldsymbol{x}_j, \boldsymbol{x}_l),\, \omega \in \mathbb{R},\, j, l = 1, \ldots, n\big\}, \tag{4.5}$$

with

$$\hat{v}(\omega, \boldsymbol{x}_j, \boldsymbol{x}_l) = \hat{u}(\omega, \boldsymbol{x}_j, \boldsymbol{x}_l) - \hat{G}_0(\omega, \boldsymbol{x}_j, \boldsymbol{x}_l)\hat{f}(\omega). \tag{4.6}$$

The incident field is given in terms of the homogeneous Green's function $\hat{G}_0(\omega, \boldsymbol{x}, \boldsymbol{y})$ defined by

$$\hat{G}_0(\omega, \boldsymbol{x}, \boldsymbol{y}) = \frac{1}{4\pi|\boldsymbol{x} - \boldsymbol{y}|} \exp\Big(i\frac{\omega}{c_o}|\boldsymbol{x} - \boldsymbol{y}|\Big), \tag{4.7}$$

that is the solution to

$$\Delta_{\boldsymbol{x}}\hat{G}_0(\omega, \boldsymbol{x}, \boldsymbol{y}) + \frac{\omega^2}{c_o^2}\hat{G}_0(\omega, \boldsymbol{x}, \boldsymbol{y}) = -\delta(\boldsymbol{x} - \boldsymbol{y}), \tag{4.8}$$

with the Sommerfeld radiation condition.

4.2 A model for the scattering medium

In this section we introduce a model for the inhomogeneous medium. We assume that the propagation speed of the medium has a homogeneous background speed value c_o and small fluctuations responsible for scattering:

$$\frac{1}{c_{\text{clu}}^2(\boldsymbol{x})} = \frac{1}{c_o^2}\big[1 + V_{\text{clu}}(\boldsymbol{x})\big], \tag{4.9}$$

where $V_{\text{clu}}(\boldsymbol{x})$ is a random process with mean zero and covariance function of the form

$$\mathbb{E}\big[V_{\mathrm{clu}}(\boldsymbol{x})V_{\mathrm{clu}}(\boldsymbol{x}')\big] = \sigma_{\mathrm{clu}}^2\sqrt{K_{\mathrm{clu}}(\boldsymbol{x})K_{\mathrm{clu}}(\boldsymbol{x}')}\,\Gamma_{\mathrm{clu}}\Big(\frac{\boldsymbol{x}-\boldsymbol{x}'}{l_{\mathrm{clu}}}\Big). \tag{4.10}$$

Here $\mathbb{E}$ stands for the expectation with respect to the distribution of the randomly scattering medium. σ_{clu} is the standard deviation of the fluctuations. The function $\boldsymbol{x} \to K_{\mathrm{clu}}(\boldsymbol{x})$ is nonnegative valued and compactly supported, it characterizes the spatial support of the scatterers (and the typical amplitude of $K_{\mathrm{clu}}(\boldsymbol{x})$ is of order one). The function $\boldsymbol{x} \to \Gamma_{\mathrm{clu}}(\boldsymbol{x}/l_{\mathrm{clu}})$ is the local correlation function. It is normalized so that $\Gamma_{\mathrm{clu}}(\boldsymbol{0}) = 1$ and $\int \Gamma_{\mathrm{clu}}(\boldsymbol{x})\mathrm{d}\boldsymbol{x} = 1$. Therefore l_{clu} can be considered as the correlation length of the random medium. We assume that the standard deviation σ_{clu} is small (smaller than one) and that correlation length l_{clu} is small (smaller than the wavelength).

For simulation purposes, the random medium fluctuations can be seen as $V_{\mathrm{clu}}(\boldsymbol{x}) = \sigma_{\mathrm{clu}}\sqrt{K_{\mathrm{clu}}(\boldsymbol{x})}Z_{\mathrm{clu}}(\boldsymbol{x}/l_{\mathrm{clu}})$ where $Z_{\mathrm{clu}}(\boldsymbol{x})$ is a stationary random process with mean zero, variance one, and correlation length equal to one. Realizations of a stationary random process can be generated by spectral methods easily.

The clutter Green's function $\hat{G}_{\mathrm{clu}}$, that is to say, the Green's function of the medium with clutter noise, is the fundamental solution of

$$\Delta_{\boldsymbol{x}}\hat{G}_{\mathrm{clu}}(\omega,\boldsymbol{x},\boldsymbol{y}) + \frac{\omega^2}{c_{\mathrm{clu}}^2(\boldsymbol{x})}\hat{G}_{\mathrm{clu}}(\omega,\boldsymbol{x},\boldsymbol{y}) = -\delta(\boldsymbol{x}-\boldsymbol{y}), \tag{4.11}$$

with the Sommerfeld radiation condition, where $c_{\mathrm{clu}}(\boldsymbol{x})$ is given by (4.9). The Lippmann-Schwinger integral equation for the clutter Green's function $\hat{G}_{\mathrm{clu}}$ defined by (4.11) is

$$\hat{G}_{\mathrm{clu}}(\omega,\boldsymbol{x},\boldsymbol{y}) = \hat{G}_0(\omega,\boldsymbol{x},\boldsymbol{y}) + \frac{\omega^2}{c_o^2}\int \hat{G}_0(\omega,\boldsymbol{x},\boldsymbol{z})V_{\mathrm{clu}}(\boldsymbol{z})\hat{G}_{\mathrm{clu}}(\omega,\boldsymbol{z},\boldsymbol{y})\mathrm{d}\boldsymbol{z}, \tag{4.12}$$

where $\hat{G}_0$ is the Green's function (4.8) of the homogeneous background medium, and $V_{\mathrm{clu}}(\boldsymbol{x})$ is the random process modeling the background fluctuations as described by (4.9).

4.3 The forward problem

4.3.1 Data set in the single-scattering regime

In this section we use the Born or single-scattering approximation for the clutter Green's function solution of (4.12) by replacing $\hat{G}_{\mathrm{clu}}$ by $\hat{G}_0$ on the right-hand side. This approximation takes into account single-scattering events for the interaction of the waves with the cluttered medium:

$$\hat{G}_{\mathrm{clu}}(\omega,\boldsymbol{x},\boldsymbol{y}) = \hat{G}_0(\omega,\boldsymbol{x},\boldsymbol{y}) + \hat{G}_1(\omega,\boldsymbol{x},\boldsymbol{y}), \tag{4.13}$$

where $\hat{G}_1$ is given by

$$\hat{G}_1(\omega, \boldsymbol{x}, \boldsymbol{y}) = \frac{\omega^2}{c_o^2} \int \hat{G}_0(\omega, \boldsymbol{x}, \boldsymbol{z}) V_{\text{clu}}(\boldsymbol{z}) \hat{G}_0(\omega, \boldsymbol{z}, \boldsymbol{y}) \mathrm{d}\boldsymbol{z}, \tag{4.14}$$

and the error is formally of order $O(\sigma_{\text{clu}}^2)$ where σ_{clu} is the standard deviation of $V_{\text{clu}}(\boldsymbol{x})$.

Let us now assume that a reflector is embedded at $\boldsymbol{z}_{\text{ref}}$ in the cluttered medium. We model the reflector by a local variation $V_{\text{ref}}(\boldsymbol{x})$ of the propagation speed as in (4.1). The full Green's function $\hat{G}$, that is to say, the Green's function of the medium with clutter in the presence of the reflector at $\boldsymbol{z}_{\text{ref}}$, is solution of

$$\Delta_{\boldsymbol{x}} \hat{G}(\omega, \boldsymbol{x}, \boldsymbol{y}) + \frac{\omega^2}{c^2(\boldsymbol{x})} \hat{G}(\omega, \boldsymbol{x}, \boldsymbol{y}) = -\delta(\boldsymbol{x} - \boldsymbol{y}). \tag{4.15}$$

The Lippmann-Schwinger integral equation for the full Green's function $\hat{G}$ is

$$\begin{aligned} \hat{G}(\omega, \boldsymbol{x}, \boldsymbol{y}) =& \hat{G}_0(\omega, \boldsymbol{x}, \boldsymbol{y}) \\ &+ \frac{\omega^2}{c_o^2} \int \hat{G}_0(\omega, \boldsymbol{x}, \boldsymbol{z}) \big(V_{\text{clu}}(\boldsymbol{z}) + V_{\text{ref}}(\boldsymbol{z})\big) \hat{G}(\omega, \boldsymbol{z}, \boldsymbol{y}) \mathrm{d}\boldsymbol{z}. \end{aligned} \tag{4.16}$$

Using again the Born approximation, we get

$$\begin{aligned} \hat{G}(\omega, \boldsymbol{x}, \boldsymbol{y}) =& \hat{G}_0(\omega, \boldsymbol{x}, \boldsymbol{y}) \\ &+ \frac{\omega^2}{c_o^2} \int \hat{G}_0(\omega, \boldsymbol{x}, \boldsymbol{z}) \big(V_{\text{clu}}(\boldsymbol{z}) + V_{\text{ref}}(\boldsymbol{z})\big) \hat{G}_0(\omega, \boldsymbol{z}, \boldsymbol{y}) \mathrm{d}\boldsymbol{z}. \end{aligned} \tag{4.17}$$

Therefore the full Green's function can be written as the sum

$$\hat{G}(\omega, \boldsymbol{x}, \boldsymbol{y}) = \hat{G}_0(\omega, \boldsymbol{x}, \boldsymbol{y}) + \hat{G}_1(\omega, \boldsymbol{x}, \boldsymbol{y}) + \hat{G}_{\text{ref}}(\omega, \boldsymbol{x}, \boldsymbol{y}). \tag{4.18}$$

The term $\hat{G}_1$ is the term in the data set that will be responsible to speckle noise in the image and it is of the form (4.14). The term $\hat{G}_{\text{ref}}$ is the term in the data set that corresponds to the reflector:

$$\hat{G}_{\text{ref}}(\omega, \boldsymbol{x}, \boldsymbol{y}) = \frac{\omega^2}{c_o^2} \int \hat{G}_0(\omega, \boldsymbol{x}, \boldsymbol{z}) V_{\text{ref}}(\boldsymbol{z}) \hat{G}_0(\omega, \boldsymbol{z}, \boldsymbol{y}) \mathrm{d}\boldsymbol{z}.$$

The approximation (4.18) is formally valid if the correction $\hat{G}_{\text{ref}}$ is small compared to $\hat{G}_0$, i.e., in the regime in which $\sigma_{\text{ref}} \ll 1$, with an error that is formally of order $O(\sigma_{\text{ref}}^2)$. We also assume that the diameter l_{ref} of the scattering region Ω_{ref} is small compared to the typical wavelength. We can then model the reflector by a point reflector (the point interaction approximation)

$$V_{\text{ref}}(\boldsymbol{x}) \approx \sigma_{\text{ref}} l_{\text{ref}}^3 \delta(\boldsymbol{x} - \boldsymbol{z}_{\text{ref}}), \tag{4.19}$$

and we can write the correction in the form

$$\hat{G}_{\text{ref}}(\omega, \boldsymbol{x}, \boldsymbol{y}) = \frac{\omega^2}{c_o^2} \sigma_{\text{ref}} l_{\text{ref}}^3 \hat{G}_0(\omega, \boldsymbol{x}, \boldsymbol{z}_{\text{ref}}) \hat{G}_0(\omega, \boldsymbol{z}_{\text{ref}}, \boldsymbol{y}). \tag{4.20}$$

The data set is therefore of the form

$$\hat{v}(\omega, \boldsymbol{x}_j, \boldsymbol{x}_l) = \hat{v}_{\text{ref}}(\omega, \boldsymbol{x}_j, \boldsymbol{x}_l) + \hat{w}(\omega, \boldsymbol{x}_j, \boldsymbol{x}_l), \tag{4.21}$$

for $j, l = 1, \ldots, n$, with $\hat{v}_{\text{ref}}$ the ideal data set in the absence of random fluctuations of the medium

$$\begin{aligned}\hat{v}_{\text{ref}}(\omega, \boldsymbol{x}_j, \boldsymbol{x}_l) =& \hat{G}_{\text{ref}}\big(\omega, \boldsymbol{x}_j, \boldsymbol{x}_l\big)\hat{f}(\omega) \\ =& \frac{\omega^2}{c_o^2}\hat{f}(\omega)\sigma_{\text{ref}} l_{\text{ref}}^3 \hat{G}_0\big(\omega, \boldsymbol{x}_j, \boldsymbol{z}_{\text{ref}}\big)\hat{G}_0\big(\omega, \boldsymbol{z}_{\text{ref}}, \boldsymbol{x}_l\big),\end{aligned} \tag{4.22}$$

and $\hat{w}$ the noise due to the interaction of the wave with the random fluctuations of the medium

$$\begin{aligned}\hat{w}(\omega, \boldsymbol{x}_j, \boldsymbol{x}_l) =& \hat{G}_1\big(\omega, \boldsymbol{x}_j, \boldsymbol{x}_l\big)\hat{f}(\omega) \\ =& \frac{\omega^2}{c_o^2}\hat{f}(\omega)\int \hat{G}_0\big(\omega, \boldsymbol{x}_j, \boldsymbol{z}\big)V_{\text{clu}}(\boldsymbol{z})\hat{G}_0\big(\omega, \boldsymbol{z}, \boldsymbol{x}_l\big)\mathrm{d}\boldsymbol{z}.\end{aligned} \tag{4.23}$$

Since the correlation length l_{clu} is small, the field $\hat{w}$ is Gaussian distributed by the Central Limit Theorem.

4.3.2 Data set in the random paraxial regime

The clutter noise model used in the previous section is rather simple and holds in a single-scattering regime, that is to say, when the medium fluctuations are very weak. When scattering is stronger another model should be used. Here we address the white-noise paraxial model. This model can be used when the central wavelength λ_0 is much smaller than the correlation length l_{clu} of the medium and the size a of the array, which are both smaller than the typical propagation distance from the array to the reflector. In this model backscattering is negligible but the wave becomes progressively incoherent as it propagates deep in the random medium. The propagation distance beyond which the wave becomes incoherent is called scattering mean free path and it will be identified below.

We assume that the array lies in the plane $z = 0$ and the medium beyond the plane $z = L$ (where the reflector to be imaged at $\boldsymbol{z}_{\text{ref}}$ is) is randomly scattering (see Figure 4.2). The data set is therefore of the form

$$\hat{v}(\omega, \boldsymbol{x}_j, \boldsymbol{x}_l) = \hat{G}_{\text{ref}}(\omega, \boldsymbol{x}_j, \boldsymbol{x}_l)\hat{f}(\omega), \tag{4.24}$$

for $j, l = 1, \ldots, n$, with

$$\hat{G}_{\text{ref}}(\omega, \boldsymbol{x}, \boldsymbol{y}) = \frac{\omega^2}{c_o^2}\sigma_{\text{ref}} l_{\text{ref}}^3 \hat{G}_{\text{clu}}(\omega, \boldsymbol{x}, \boldsymbol{z}_{\text{ref}})\hat{G}_{\text{clu}}(\omega, \boldsymbol{z}_{\text{ref}}, \boldsymbol{y}), \tag{4.25}$$

and $\hat{G}_{\text{clu}}$ is the Green's function of the random medium. If $\boldsymbol{y} = (\boldsymbol{y}_\perp, 0)$ and we denote $\boldsymbol{z}_{\text{ref}} = (\boldsymbol{z}_{\text{r}_\perp}, z_{\text{r}})$, it is of the form

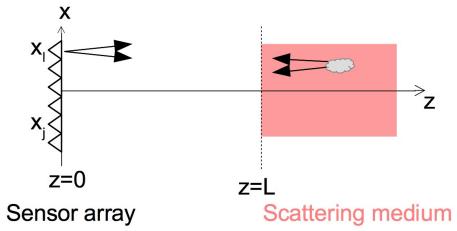


Fig. 4.2. Reverse-Time imaging set-up in the random paraxial regime. $\boldsymbol{x}_l$ emits a short pulse that propagates through the medium, is backscattered by the reflector, and is recorded by the sensor at $\boldsymbol{x}_j$. The array is in the plane $z = 0$, the medium is randomly scattering in the half space $z > L$. The reflector to be imaged in embedded in the random medium.

$$\hat{G}_{\mathrm{clu}}(\omega, \boldsymbol{z}_{\mathrm{ref}}, \boldsymbol{y}) = \frac{ic_o}{2\omega} \exp\left(i\frac{\omega}{c_o} z_{\mathrm{r}}\right) \hat{g}(\omega, \boldsymbol{z}_{\mathrm{r}_\perp}, z_{\mathrm{r}}, \boldsymbol{y}_\perp), \tag{4.26}$$

where $(\boldsymbol{x}_\perp, z) \to \hat{g}(\omega, \boldsymbol{x}_\perp, z, \boldsymbol{y}_\perp)$ is solution of the Itô-Schrödinger equation

$$\mathrm{d}\hat{g} = \frac{ic_o}{2\omega} \Delta_{\boldsymbol{x}_\perp} \hat{g} \mathrm{d}z + \frac{i\omega}{2c_o} \hat{g} \circ \mathrm{d}B(\boldsymbol{x}_\perp, z), \tag{4.27}$$

starting from $\hat{g}(\omega, \boldsymbol{x}_\perp, z = 0, \boldsymbol{y}_\perp) = \delta(\boldsymbol{x}_\perp - \boldsymbol{y}_\perp)$. Here the symbol $\circ$ stands for the Stratonovich stochastic integral and $B(\boldsymbol{x}_\perp, z)$ is a real-valued Brownian field over $\mathbb{R}^2 \times [L, \infty)$ with covariance

$$\mathbb{E}[B(\boldsymbol{x}_\perp, z) B(\boldsymbol{x}'_\perp, z')] = \begin{cases} \min\{z - L, z' - L\} C(\boldsymbol{x}_\perp - \boldsymbol{x}'_\perp) & \text{if } z, z' > L, \\ 0 & \text{otherwise.} \end{cases} \tag{4.28}$$

The model (4.27) can be obtained from the scalar wave equation by a separation of scales technique in which the three-dimensional fluctuations of the speed of propagation $c^2(\boldsymbol{x}_\perp, z)$ are described by a zero-mean random process $V_{\mathrm{clu}}(\boldsymbol{x}_\perp, z)$ that is stationary in the half space $z > L$ and that has strong mixing properties:

$$\frac{1}{c^2_{\mathrm{clu}}(\boldsymbol{x}_\perp, z)} = \frac{1}{c_o^2} \left[1 + V_{\mathrm{clu}}(\boldsymbol{x}_\perp, z)\right].$$

If the covariance function of the process V_{clu} is of the form (4.10) with $K_{\mathrm{clu}}((\boldsymbol{x}_\perp, z)) = \mathbf{1}_{[L,\infty)}(z)$, then the covariance function $C(\boldsymbol{x}_\perp)$ in (4.28) is given by [23]:

$$C(\boldsymbol{x}_\perp) = \sigma^2_{\mathrm{clu}} l_{\mathrm{clu}} \mathcal{C}\left(\frac{\boldsymbol{x}_\perp}{l_{\mathrm{clu}}}\right), \qquad \mathcal{C}(\boldsymbol{X}_\perp) = \int_{-\infty}^{\infty} \Gamma_{\mathrm{clu}}(\boldsymbol{X}_\perp, z) \mathrm{d}z. \tag{4.29}$$

When the medium is homogeneous, the Green's function of the medium is equal to the homogeneous Green's function in the paraxial regime (or Fresnel approximation):

$$\hat{G}_0(\omega, \boldsymbol{x}, \boldsymbol{z}_{\rm ref}) = \frac{1}{4\pi z_{\rm r}} \exp\Big(i\frac{\omega}{c_o} z_{\rm r} + i\frac{\omega}{c_o}\frac{|\boldsymbol{x}_\perp - \boldsymbol{z}_{{\rm r}\perp}|^2}{2z_{\rm r}}\Big). \tag{4.30}$$

When the medium is scattering, the moments of the Green's function of the random medium can be computed by Itô's formula. In particular, for any $\boldsymbol{x} = (\boldsymbol{x}_\perp, 0)$,

$$\mathbb{E}\big[\hat{G}_{\rm clu}(\omega, \boldsymbol{x}, \boldsymbol{z}_{\rm ref})\big] = \hat{G}_0(\omega, \boldsymbol{x}, \boldsymbol{z}_{\rm ref}) \exp\Big(-\frac{\sigma_{\rm clu}^2 l_{\rm clu} \mathcal{C}(\boldsymbol{0})\omega^2}{8c_o^2}(z_{\rm r} - L)\Big),$$

which shows that the scattering mean free path at frequency ω in the random half-space $z \in (L, \infty)$ is

$$\ell_{\rm sca}(\omega) = \frac{8c_o^2}{\sigma_{\rm clu}^2 l_{\rm clu} \mathcal{C}(\boldsymbol{0})\omega^2}.$$

Moreover, for any $\boldsymbol{x}, \boldsymbol{x}'$ in the plane $z = 0$, the second-order moments are [23, 24]

$$\begin{aligned}\mathbb{E}\big[\hat{G}_{\rm clu}(\omega, \boldsymbol{x}, \boldsymbol{z}_{\rm ref})\hat{G}_{\rm clu}(\omega, \boldsymbol{x}', \boldsymbol{z}_{\rm ref})\big] =& \hat{G}_0(\omega, \boldsymbol{x}, \boldsymbol{z}_{\rm ref})\hat{G}_0(\omega, \boldsymbol{x}', \boldsymbol{z}_{\rm ref}) \\ &\times \exp\Big(-\frac{\sigma_{\rm clu}^2 l_{\rm clu} \mathcal{C}(\boldsymbol{0})\omega^2}{4c_o^2}(z_{\rm r} - L)\Big),\end{aligned} \tag{4.31}$$

$$\begin{aligned}\mathbb{E}\big[\hat{G}_{\rm clu}(\omega, \boldsymbol{x}, \boldsymbol{z}_{\rm ref})\overline{\hat{G}_{\rm clu}(\omega, \boldsymbol{x}', \boldsymbol{z}_{\rm ref})}\big] =& \hat{G}_0(\omega, \boldsymbol{x}, \boldsymbol{z}_{\rm ref})\overline{\hat{G}_0(\omega, \boldsymbol{x}', \boldsymbol{z}_{\rm ref})} \\ &\times \exp\Big(-\frac{|\boldsymbol{x}_\perp - \boldsymbol{x}'_\perp|^2}{X_c^2}\Big),\end{aligned} \tag{4.32}$$

with the transverse correlation radius of the field

$$X_c^2 = \frac{3c_o^2 z_{\rm r}^2 l_{\rm clu}}{\omega_0^2 \sigma_{\rm clu}^2 \mathcal{D}(z_{\rm r} - L)^3}, \tag{4.33}$$

and $\mathcal{D} = -\Delta\mathcal{C}(\boldsymbol{0})$. Here and below we neglect the dependence of X_c with respect to ω which holds when the source bandwidth B is smaller than the central frequency ω_0.

4.4 The imaging function

In the presence of a point reflector at $\boldsymbol{z}_{\rm ref}$ and in the presence of clutter noise the data set is of the form (4.21). We study the effect of the clutter noise on the reverse-time imaging function defined by

$$\mathcal{I}_{\rm RT}(\boldsymbol{x}) = \frac{1}{2\pi n^2} \sum_{j,l=1}^{n} \int {\rm d}\omega \hat{G}_0(\omega, \boldsymbol{x}, \boldsymbol{x}_j)\hat{G}_0(\omega, \boldsymbol{x}, \boldsymbol{x}_l)\overline{\hat{v}(\omega, \boldsymbol{x}_j, \boldsymbol{x}_l)}. \tag{4.34}$$

4.4.1 The imaging function without clutter noise

In the absence of noise $\sigma_{\rm clu} = 0$ the imaging function is $\mathcal{I}_{\rm RT}(\boldsymbol{x}) = \mathcal{I}_{\rm RT,0}(\boldsymbol{x})$ where $\mathcal{I}_{\rm RT,0}(\boldsymbol{x})$ is given by

$$\mathcal{I}_{\rm RT,0}(\boldsymbol{x}) = \sigma_{\rm ref} l_{\rm ref}^3 \mathcal{H}(\boldsymbol{x}, \boldsymbol{z}_{\rm ref}), \tag{4.35}$$

with

$$\mathcal{H}(\boldsymbol{x}, \boldsymbol{y}) = \frac{1}{2\pi} \int \mathrm{d}\omega \frac{\omega^2}{c_o^2} \hat{f}(\omega) \hat{h}(\omega, \boldsymbol{x}, \boldsymbol{y}), \tag{4.36}$$

$$\hat{h}(\omega, \boldsymbol{x}, \boldsymbol{y}) = \Big(\frac{1}{n} \sum_{j=1}^{n} \overline{\hat{G}_0(\omega, \boldsymbol{x}, \boldsymbol{x}_j)} \hat{G}_0(\omega, \boldsymbol{y}, \boldsymbol{x}_j) \Big)^2. \tag{4.37}$$

The function $\boldsymbol{x} \to \mathcal{H}(\boldsymbol{x}, \boldsymbol{z}_{\rm ref})$ is the point spread function that describes the spatial profile of the peak obtained at the reflector location $\boldsymbol{z}_{\rm ref}$ in the imaging function when the reflector is point-like.

Full-aperture array. If the sensor array is dense (i.e. the inter-sensor distance is smaller than half a wavelength) and completely surrounds the region of interest (for instance, it covers the surface of the ball with center at $\mathbf{0}$ and radius L) then Helmholtz-Kirchhoff theorem states that $\hat{h}(\omega, \boldsymbol{x}, \boldsymbol{y})$ is proportional to the square of the imaginary part of the Green's function $\hat{G}_0(\omega, \boldsymbol{x}, \boldsymbol{y})$. We find

$$\hat{h}(\omega, \boldsymbol{x}, \boldsymbol{y}) = C_L \widetilde{h}(\omega, \boldsymbol{x} - \boldsymbol{y}), \text{ where } \widetilde{h}(\omega, \boldsymbol{x}) = \mathrm{sinc}^2\Big(\frac{\omega |\boldsymbol{x}|}{c_o} \Big), \tag{4.38}$$

and $C_L = 1/(4\pi L)^4$. Therefore, when the bandwidth B of f is smaller than its central frequency ω_0, for instance, when the source is a modulated Gaussian with central frequency ω_0 and bandwidth $B \ll \omega_0$:

$$f(t) = \cos(\omega_0 t) \exp\big(-\frac{B^2 t^2}{2} \big), \tag{4.39}$$

$$\hat{f}(\omega) = \frac{\sqrt{\pi}}{\sqrt{2}B} \exp\Big(-\frac{(\omega - \omega_0)^2}{2B^2} \Big) + \frac{\sqrt{\pi}}{\sqrt{2}B} \exp\Big(-\frac{(\omega + \omega_0)^2}{2B^2} \Big), \tag{4.40}$$

we have

$$\mathcal{H}(\boldsymbol{x}, \boldsymbol{z}_{\rm ref}) = C_L \frac{\omega_0^2}{c_o^2} \widetilde{\mathcal{H}}(\boldsymbol{x} - \boldsymbol{z}_{\rm ref}), \tag{4.41}$$

where

$$\widetilde{\mathcal{H}}(\boldsymbol{x}) = \mathrm{sinc}^2\Big(\frac{2\pi |\boldsymbol{x}|}{\lambda_0} \Big), \tag{4.42}$$

and $\lambda_0 = 2\pi c_o/\omega_0$ is the central wavelength. This shows that the width of the point spread function $\widetilde{\mathcal{H}}(\boldsymbol{x})$ is of the order of $\lambda_0/2$, which is the Abbe diffraction limit [**?**]. Note also that the function $\widetilde{\mathcal{H}}(\boldsymbol{x})$ decays slowly, as $1/|\boldsymbol{x}|^2$, which

will turn out to be problematic when addressing scattering media.

Finite-aperture array. If the sensor array is dense and occupies the domain $D_a \times \{0\}$, with $D_a \subset \mathbb{R}^2$ with diameter a, and the search region is a domain Ω around $(0,0,L)$, then in the Fresnel diffraction regime $\lambda_0 \ll a \ll L$ with $\lambda_0^{1/2} L^{3/2} \gg a^2 \gg \lambda_0 L$ we have

$$\hat{h}(\omega, \boldsymbol{x}, \boldsymbol{y}) = C_L \widetilde{h}(\omega, \boldsymbol{x} - \boldsymbol{y}), \tag{4.43}$$

where, for $\boldsymbol{x} = (\boldsymbol{x}_\perp, x_3)$,

$$\widetilde{h}(\omega, \boldsymbol{x}) = e^{-2i\frac{\omega}{c_o}x_3} \left(\frac{1}{|D_a|} \int_{D_a} \exp\Big(-i \frac{\omega \boldsymbol{y}_\perp}{c_o L} \cdot \boldsymbol{x}_\perp - i \frac{\omega |\boldsymbol{y}_\perp|^2}{2 c_o L^2} x_3 \Big) \mathrm{d}\boldsymbol{y}_\perp \right)^2. \tag{4.44}$$

This shows that the width of the function $\widetilde{h}(\omega, \boldsymbol{x})$ is of the order of $\lambda L/a$ in the transverse directions $(\boldsymbol{x}_\perp)$ and $\lambda L^2/a^2$ in the longitudinal direction (x_3), where $\lambda = 2\pi c_o/\omega$ is the wavelength associated to the frequency ω. These are the classical Rayleigh resolution formulas for time-harmonic waves [**?**, Sections 8.5 and 8.8].

If the bandwidth B of the pulse is smaller than the central frequency ω_0 and such that $B \ll \omega_0 a^2/L^2$, for instance, when the source is a modulated Gaussian (4.39) with central frequency ω_0 and bandwidth B, then the range resolution is the same one as in the time-harmonic regime and we have

$$\mathcal{H}(\boldsymbol{x}, \boldsymbol{z}_{\mathrm{ref}}) = C_L \frac{\omega_0^2}{c_o^2} \widetilde{\mathcal{H}}(\boldsymbol{x} - \boldsymbol{z}_{\mathrm{ref}}), \tag{4.45}$$

where

$$\widetilde{\mathcal{H}}(\boldsymbol{x}) = \frac{1}{2} \exp\Big(-i \frac{2\omega_0 x_3}{c_o}\Big) \left(\frac{1}{|D_a|} \int_{D_a} \exp\Big(-i \frac{2\pi \boldsymbol{y}_\perp}{\lambda_0 L} \cdot \boldsymbol{x}_\perp - i \frac{\pi |\boldsymbol{y}_\perp|^2}{\lambda_0 L^2} x_3\Big) \mathrm{d}\boldsymbol{y}_\perp \right)^2 + cc. \tag{4.46}$$

This shows that the width of the point spread function $\widetilde{\mathcal{H}}(\boldsymbol{x})$ is of the order of $\lambda_0 L/a$ in the transverse directions $(\boldsymbol{x}_\perp)$ and $\lambda_0 L^2/a^2$ in the longitudinal direction (x_3). Note the loss of resolution compared to the full-aperture case, by a factor L/a in the transverse directions and $(L/a)^2$ in the longitudinal direction. Note also that the function $\widetilde{\mathcal{H}}(\boldsymbol{x})$ decays slowly, as $1/|\boldsymbol{x}_\perp|^2$ in the transverse direction and as $1/|x_3|$ in the longitudinal direction. More precisely, if the array is square with sidelength a, i.e.

$$D_a = [-a/2, a/2]^2, \tag{4.47}$$

then

$$\widetilde{\mathcal{H}}(\boldsymbol{x}) = \frac{1}{2} \exp\Big(-i \frac{2\omega_0 x_3}{c_o} \Big) \Psi\Big(\frac{a}{\lambda_0 L} \boldsymbol{x}_\perp, \frac{a^2}{\lambda_0 L^2} x_3 \Big) + cc, \tag{4.48}$$

with

$$\Psi(x_1,x_2,x_3)=\left(\int_{[-1/2,1/2]^2}\exp\Big(-i2\pi(s_1x_1+s_2x_2-i\pi(s_1^2+s_2^2)x_3\Big)\mathrm{d}s_1\mathrm{d}s_2\right)^2. \tag{4.49}$$

If the bandwidth B of the pulse is smaller than the central frequency ω_0 but such that $B \gg \omega_0 a^2/L^2$, for instance, when the source is a modulated Gaussian (4.39) with central frequency ω_0 and bandwidth B, then the range resolution is bandwidth-limited and we have

$$\mathcal{H}(\boldsymbol{x},\boldsymbol{z}_{\mathrm{ref}})=C_L\frac{\omega_0^2}{c_o^2}\widetilde{\mathcal{H}}(\boldsymbol{x}-\boldsymbol{z}_{\mathrm{ref}}), \tag{4.50}$$

where

$$\begin{aligned}\widetilde{\mathcal{H}}(\boldsymbol{x}) =&\frac{1}{2}\exp\Big(-i\frac{2\omega_0x_3}{c_o}\Big)\exp\Big(-2\frac{B^2x_3^2}{c_o^2}\Big)\\ &\times\left(\frac{1}{|D_a|}\int_{D_a}\exp\Big(-i\frac{2\pi\boldsymbol{y}_\perp}{\lambda_0L}\cdot\boldsymbol{x}_\perp\Big)\mathrm{d}\boldsymbol{y}_\perp\right)^2+cc.\end{aligned} \tag{4.51}$$

If the array is square with sidelength a, i.e.

$$D_a=[-a/2,a/2]^2, \tag{4.52}$$

then

$$\widetilde{\mathcal{H}}(\boldsymbol{x})=\cos\Big(\frac{2\omega_0x_3}{c_o}\Big)\exp\Big(-2\frac{B^2x_3^2}{c_o^2}\Big)\mathrm{sinc}^2\Big(\frac{\pi a}{\lambda_0L}x_1\Big)\mathrm{sinc}^2\Big(\frac{\pi a}{\lambda_0L}x_2\Big). \tag{4.53}$$

This shows that the width of the point spread function $\widetilde{\mathcal{H}}(\boldsymbol{x})$ is of the order of λ_0L/a in the transverse directions ($\boldsymbol{x}_\perp$) and $c_o/(2B)$ in the longitudinal direction (x_3). Note the loss of resolution compared to the full-aperture case, by a factor L/a in the transverse directions and ω_0/B in the longitudinal direction. Note also that the function $\widetilde{\mathcal{H}}(\boldsymbol{x})$ decays slowly, as $1/|\boldsymbol{x}_\perp|^2$ in the transverse direction, but may decay fast in the longitudinal direction depending on the source spectrum (here, a Gaussian).

4.4.2 The imaging function with clutter noise in the single scattering regime

In the presence of clutter noise in the single-scattering regime the imaging function is a real Gaussian random field. Its mean is the unperturbed imaging function $\mathcal{I}_{\mathrm{RT},0}$ defined by (4.35) and the covariance function of the imaging function is:

$$\begin{aligned}\mathrm{Cov}\big(\mathcal{I}_{\mathrm{RT}}(\boldsymbol{x}),\mathcal{I}_{\mathrm{RT}}(\boldsymbol{x}')\big)=\sigma_{\mathrm{clu}}^2\iint\mathrm{d}\boldsymbol{y}\mathrm{d}\boldsymbol{y}'\mathcal{H}(\boldsymbol{x},\boldsymbol{y})\mathcal{H}(\boldsymbol{y}',\boldsymbol{x}')\\ \times\sqrt{K_{\mathrm{clu}}(\boldsymbol{y})K_{\mathrm{clu}}(\boldsymbol{y}')}\Gamma_{\mathrm{clu}}\Big(\frac{\boldsymbol{y}-\boldsymbol{y}'}{l_{\mathrm{clu}}}\Big),\end{aligned} \tag{4.54}$$

which follows from the expression (4.23) of the cluttered noise and the form (4.10) of the covariance function of the medium fluctuations.

Using the fact that the correlation length of the medium is small, this can be simplified as

$$\mathrm{Cov}\big(\mathcal{I}_{\mathrm{RT}}(\boldsymbol{x}),\mathcal{I}_{\mathrm{RT}}(\boldsymbol{x}')\big) = \sigma_{\mathrm{clu}}^2 l_{\mathrm{clu}}^3 \int \mathrm{d}\boldsymbol{y}\mathcal{H}(\boldsymbol{x},\boldsymbol{y})K_{\mathrm{clu}}(\boldsymbol{y})\mathcal{H}(\boldsymbol{y},\boldsymbol{x}'). \quad (4.55)$$

If we assume that the random scatterers are uniformly distributed in the search region, i.e. $K_{\mathrm{clu}} \equiv 1$, and that $\mathcal{H}(\boldsymbol{x},\boldsymbol{y}) = C_L(\omega_0^2/c_o^2)\widetilde{\mathcal{H}}(\boldsymbol{x}-\boldsymbol{y})$ as in the case of a full-aperture array or finite-aperture array discussed above, then we have

$$\mathrm{Cov}\big(\mathcal{I}_{\mathrm{RT}}(\boldsymbol{x}),\mathcal{I}_{\mathrm{RT}}(\boldsymbol{x}')\big) = \sigma_{\mathrm{clu}}^2 l_{\mathrm{clu}}^3 C_L^2 \frac{\omega_0^4}{c_o^4}\widetilde{\mathcal{H}} * \widetilde{\mathcal{H}}(\boldsymbol{x}-\boldsymbol{x}'), \quad (4.56)$$

where the star stands for the convolution in $\mathbb{R}^3$. This shows that the speckle noise $\mathcal{I}_{\mathrm{RT}} - \mathcal{I}_{\mathrm{RT},0}$ in the image is a stationary Gaussian random field with mean zero, variance

$$\sigma_{\mathrm{noise}}^2 = \sigma_{\mathrm{clu}}^2 l_{\mathrm{clu}}^3 C_L^2 \frac{\omega_0^4}{c_o^4}\widetilde{\mathcal{H}} * \widetilde{\mathcal{H}}(\boldsymbol{0}), \quad (4.57)$$

and covariance function:

$$\mathrm{Cov}\big(\mathcal{I}_{\mathrm{RT}}(\boldsymbol{x}),\mathcal{I}_{\mathrm{RT}}(\boldsymbol{x}')\big) = \sigma_{\mathrm{noise}}^2 \mathcal{H}_{\mathrm{noise}}(\boldsymbol{x}-\boldsymbol{x}'), \quad (4.58)$$

with

$$\mathcal{H}_{\mathrm{noise}}(\boldsymbol{x}) = \frac{\widetilde{\mathcal{H}} * \widetilde{\mathcal{H}}(\boldsymbol{x})}{\widetilde{\mathcal{H}} * \widetilde{\mathcal{H}}(\boldsymbol{0})}. \quad (4.59)$$

This random field is a speckle pattern whose hotspot profiles are close to the function $\mathcal{H}_{\mathrm{noise}}(\boldsymbol{x})$, which is (proportional to) the autoconvolution of the point spread function of the imaging function. Here the hotspot profile refers to the local shape of the field around a local maximum (see Appendix 6.5).

Note also that, when the random scatterers are not uniformly distributed in the search region, i.e. K_{clu} is not constant, then the slow decay (as a power law) of the function $\widetilde{\mathcal{H}}$ implies that the random scatterers in a far region can generate speckle noise everywhere in the image as shown by (4.55). As a consequence it is very difficult to image through a scattering layer, even if it is thin and the reflector to be imaged is beyond it and in a -relatively- homogeneous region. This is a serious drawback for reverse-time imaging.

Full-aperture array. In the case in which the array completely surrounds the region of interest, $\mathcal{I}_{\mathrm{RT}}(\boldsymbol{x})$ is a Gaussian random field with mean $\mathcal{I}_{\mathrm{RT},0}(\boldsymbol{x})$ given by (4.35), variance

$$\sigma^2_{\text{noise}} = \sigma^2_{\text{clu}} l^3_{\text{clu}} C_L^2 \frac{\omega_0^4}{c_o^4} \frac{\lambda_0^3}{8\pi}, \tag{4.60}$$

and covariance function:

$$\text{Cov}\big(\mathcal{I}_{\text{RT}}(\boldsymbol{x}), \mathcal{I}_{\text{RT}}(\boldsymbol{x}')\big) = \sigma^2_{\text{noise}} h_{\text{noise}}\Big(\frac{4\pi|\boldsymbol{x}-\boldsymbol{x}'|}{\lambda_0}\Big), \tag{4.61}$$

with the normalized function ($h_{\text{noise}}(0) = 1$)

$$h_{\text{noise}}(x) = \frac{\text{Si}(x)}{x}, \qquad \text{Si}(x) = \int_0^x \text{sinc}(y)\text{d}y. \tag{4.62}$$

The speckle pattern is made of hotspots with typical radius λ_0 and typical amplitude σ_{noise}. The typical shape of the hotspot is given by the function h_{noise}, that has a slow power law decay as $1/x$. The signal to noise ratio in the image is

$$\text{SNR} = \frac{\mathcal{I}_{\text{RT},0}(\boldsymbol{z}_{\text{ref}})}{\sigma_{\text{noise}}} = \frac{2\sqrt{2\pi}\sigma_{\text{ref}} l^3_{\text{ref}}}{\sigma_{\text{clu}} l^{3/2}_{\text{clu}} \lambda_0^{3/2}}. \tag{4.63}$$

The hotspot volume is defined as

$$V_c = \frac{4\pi^2}{(\det \boldsymbol{\Lambda})^{1/2}}, \qquad \boldsymbol{\Lambda} = \Big(-\partial^2_{x_j x_l} h_{\text{noise}}\big(\frac{4\pi|\boldsymbol{x}|}{\lambda_0}\big) \,|_{\boldsymbol{x}=\boldsymbol{0}} \Big)_{j,l=1,\ldots,3}. \tag{4.64}$$

Here the matrix $\boldsymbol{\Lambda}$ is proportional to the identity and the hotspot volume is proportional to the central wavelength to the power three:

$$\boldsymbol{\Lambda} = \frac{16\pi^2}{9\lambda_0^2}\mathbf{I}, \qquad V_c = \frac{27}{16\pi}\lambda_0^3. \tag{4.65}$$

The maximum of the function $\mathcal{I}_{\text{RT}} - \mathcal{I}_{\text{RT},0}$ over a domain Ω whose volume is much larger than the hotspot volume is a random quantity described in Appendix 6.5, which is equal to the deterministic value (3.27) to leading order in $|\Omega|/V_c$, with the values of σ_{noise} and V_c as given by (4.60) and (4.65). This result shows that the SNR (4.63) should be considered with cautious. The speckle pattern may have a local maximum whose peak amplitude is much larger than σ_{noise} and that can be misinterpreted as a reflector.

Finite-aperture array. In the case in which the sensor array is dense and occupies the square domain $D_a \times \{0\}$, with $D_a = [-a/2, a/2]^2 \subset \mathbb{R}^2$, the bandwidth B of the pulse is smaller than the central frequency ω_0 but such that $B \gg \omega_0 a^2/L^2$, and the search region is a domain Ω around $(0, 0, L)$, then the field $\mathcal{I}_{\text{RT}}(\boldsymbol{x})$ is a Gaussian random field with mean $\mathcal{I}_{\text{RT},0}(\boldsymbol{x})$ given by (4.35), variance

$$\sigma^2_{\text{noise}} = \sigma^2_{\text{clu}} l^3_{\text{clu}} C_L^2 \frac{\omega_0^4}{c_o^4} \frac{c_o\sqrt{\pi}}{B} \frac{\lambda_0^2 L^2}{9a^2}, \tag{4.66}$$

and covariance function:

$$\mathrm{Cov}\big(\mathcal{I}_{\mathrm{RT}}(\boldsymbol{x}),\mathcal{I}_{\mathrm{RT}}(\boldsymbol{x}')\big) = \sigma_{\mathrm{noise}}^2 \cos\Big(2\frac{\omega_0(x_3-x_3')}{c_o}\Big) \times h_{\mathrm{noise}}\Big(\frac{2\pi a(x_1-x_1')}{\lambda_0 L},\frac{2\pi a(x_2-x_2')}{\lambda_0 L},\frac{2B(x_3-x_3')}{c_o}\Big), \quad (4.67)$$

with the normalized function ($h_{\mathrm{noise}}(\mathbf{0})=1$)

$$h_{\mathrm{noise}}(x_1,x_2,x_3) = 36\frac{1-\mathrm{sinc}(x_1)}{x_1^2}\frac{1-\mathrm{sinc}(x_2)}{x_2^2}\exp\Big(-\frac{x_3^2}{4}\Big). \quad (4.68)$$

The speckle pattern is made of hotspots with typical radius $\lambda_0 L/a$ in the transverse direction, typical radius $c_o/(2B)$ in the longitudinal direction, and typical amplitude σ_{noise}. The signal to noise ratio in the image is

$$\mathrm{SNR} = \frac{\mathcal{I}_{\mathrm{RT},0}(\boldsymbol{z}_{\mathrm{ref}})}{\sigma_{\mathrm{noise}}} = \frac{3\sqrt{2}\sqrt[4]{\pi}\sigma_{\mathrm{ref}} l_{\mathrm{ref}}^3}{\sigma_{\mathrm{clu}} l_{\mathrm{clu}}^{3/2}\lambda_0^{3/2}}\frac{B^{1/2}a}{\omega_0^{1/2}L}. \quad (4.69)$$

Note the SNR reduction compared to the full-aperture case, by a factor of the order of $(B/\omega_0)^{1/2}(a/L)$.

The hotspot volume is defined as before as $V_c = 4\pi^2(\det \boldsymbol{\Lambda})^{-1/2}$ with

$$\boldsymbol{\Lambda} = \Big(-\partial^2_{x_j x_l} h_{\mathrm{noise}}\Big(\frac{2\pi a x_1}{\lambda_0 L},\frac{2\pi a x_2}{\lambda_0 L},\frac{2Bx_3}{c_o}\Big)\,|_{\boldsymbol{x}=\mathbf{0}}\Big)_{j,l=1,\dots,3}. \quad (4.70)$$

Here $\boldsymbol{\Lambda}$ is diagonal and we have

$$\Lambda_{jj} = \frac{2\pi^2 a^2}{5\lambda_0^2 L^2}, \quad j=1,2, \qquad \Lambda_{33} = \frac{2B^2}{c_o^2}, \qquad V_c = \frac{5\lambda_0^3}{\sqrt{2}\pi}\frac{a^2\omega_0}{L^2 B}. \quad (4.71)$$

The maximum of the function over a domain Ω whose volume is much larger than the hotspot volume is given by (3.27) as before with the values of σ_{noise} and V_c as given by (4.66) and (4.71).

4.4.3 The imaging function with clutter noise in the random paraxial regime

In the presence of clutter noise in the random paraxial regime the imaging function is a real random field, but its statistics is not Gaussian. However the moments of the cluttered Green's function are known, so it is possible to carry out a detailed statistical analysis of the imaging function.

From the expression (4.31) of the second-order moments, we obtain that the mean imaging function has the form

$$\begin{aligned}\mathbb{E}\big[\mathcal{I}_{\mathrm{RT}}(\boldsymbol{x})\big] &= \mathcal{I}_{\mathrm{RT},0}(\boldsymbol{x})\exp\Big(-\frac{\sigma_{\mathrm{clu}}^2 l_{\mathrm{clu}}\mathcal{C}(\mathbf{0})\omega^2}{4c_o^2}(z_{\mathrm{r}}-L)\Big)\\ &= \sigma_{\mathrm{ref}} l_{\mathrm{ref}}^3 C_L\frac{\omega_0^2}{c_o^2}\widetilde{\mathcal{H}}(\boldsymbol{x}-\boldsymbol{z}_{\mathrm{ref}})\exp\Big(-\frac{\sigma_{\mathrm{clu}}^2 l_{\mathrm{clu}}\mathcal{C}(\mathbf{0})\omega^2}{4c_o^2}(z_{\mathrm{r}}-L)\Big),\end{aligned} \quad (4.72)$$

where $\widetilde{\mathcal{H}}$ is defined by (4.53) and $\mathcal{I}_{\mathrm{RT},0}(\boldsymbol{x})$ is the imaging function obtained when the medium is homogeneous. The peak associated to the reflector is exponentially damped as its depth increases, which is due to a loss of coherence by scattering. This is the main difference with the clutter noise in the single-scattering regime addressed in the previous section, in which the main peak is always present with the same amplitude.

Using the formulas for the fourth-order moment of the random Green's function obtained in [25], we find that the covariance function is

$$\begin{aligned}\mathrm{Cov}\big(\mathcal{I}_{\mathrm{RT}}(\boldsymbol{x}),\mathcal{I}_{\mathrm{RT}}(\boldsymbol{x}')\big) =&\sigma_{\mathrm{ref}}^2 l_{\mathrm{ref}}^6 C_L^2 \frac{\omega_0^4}{c_o^4}\exp\Big(-\frac{(Z-z_{\mathrm{r}})^2}{Z_c^2}\Big)\\ &\times\frac{2\pi^2 X_c^4}{a^4}\exp\Big(-\frac{B^2\zeta^2}{c_o^2}\Big)\cos\Big(2\frac{\omega_0}{c_o}\zeta\Big)\\ &\times\mathrm{sinc}^2\Big(\frac{\pi a(x_1-x_1')}{\lambda_0 L}\Big)\mathrm{sinc}^2\Big(\frac{\pi a(x_2-x_2')}{\lambda_0 L}\Big),\end{aligned}\tag{4.73}$$

where we have denoted $\boldsymbol{x}=(\boldsymbol{X}_\perp+\boldsymbol{\rho}_\perp/2, Z+\zeta/2)$ and $\boldsymbol{x}'=(\boldsymbol{X}_\perp-\boldsymbol{\rho}_\perp/2, Z-\zeta/2)$, we have assumed $X_c \ll a$, and we have introduced

$$Z_c=\frac{c_o}{2B}\Big(1+\frac{\sigma_{\mathrm{clu}}^2 l_{\mathrm{clu}}\mathcal{C}(\mathbf{0})(z_{\mathrm{r}}-L)B^2}{c_o^2}\Big)^{1/2}.$$

As explained below, the depth $z_{\mathrm{r}}-L$ cannot be much larger than the scattering mean free path and $B \ll \omega_0$, so that we have in fact $Z_c \simeq c_o/(2B)$. The fluctuations in the image has the form of a speckle pattern concentrated around the depth z_{r} whose hotspots have transverse radius of the order of $\lambda_0 L/a$. The variance of the fluctuations relative to the maximal amplitude of the unperturbed imaging function is $2\pi^2 X_c^4/a^4$, independently of the depth z_{r} of the reflector. Therefore the SNR of the image (when $X_c \ll a$) is

$$\mathrm{SNR}=\frac{\mathcal{I}_{\mathrm{RT},0}(\boldsymbol{z}_{\mathrm{ref}})}{\mathrm{Var}(\mathcal{I}_{\mathrm{RT}}(\boldsymbol{z}_{\mathrm{ref}}))^{1/2}}=\frac{a^2}{\sqrt{2}\pi X_c^2}\exp\Big(-\frac{\sigma_{\mathrm{clu}}^2 l_{\mathrm{clu}}\mathcal{C}(\mathbf{0})\omega_0^2}{4c_o^2}(z_{\mathrm{r}}-L)\Big).\tag{4.74}$$

From this expression, we can state the following assertions:
- A large array is necessary to get a high SNR. Therefore the size of the array is good both for resolution and stability. The SNR scales as the number of independent patches in the array, which have areas of the order of X_c^2 while the aera of the array is a^2.
- It is all the more difficult to image the reflector as it is deep. It becomes impossible when its depth is of the order of a few scattering mean free paths due to the exponential decay of the SNR.
- A large bandwidth does not improve the SNR (as long as it is smaller than the carrier frequency ω_0). This is special to the random paraxial regime that we consider in this section, in which the frequency coherence radius is rather large. Indeed, by Itô's formula,

$$\mathbb{E}\big[\hat{G}_{\mathrm{clu}}(\omega,\boldsymbol{x},\boldsymbol{z}_{\mathrm{ref}})\overline{\hat{G}_{\mathrm{clu}}(\omega',\boldsymbol{x},\boldsymbol{z}_{\mathrm{ref}})}\big] = \hat{G}_0(\omega,\boldsymbol{x},\boldsymbol{z}_{\mathrm{ref}})\overline{\hat{G}_0(\omega',\boldsymbol{x},\boldsymbol{z}_{\mathrm{ref}})} \times \exp\Big(-\frac{\sigma_{\mathrm{clu}}^2 l_{\mathrm{clu}}\mathcal{C}(\mathbf{0})(\omega-\omega')^2(z_{\mathrm{r}}-L)}{8c_o^2}\Big),$$

which shows that the frequency coherence Ω_c is given by

$$\Omega_c^2 = \frac{8c_o^2}{\sigma_{\mathrm{clu}}^2 l_{\mathrm{clu}}\mathcal{C}(\mathbf{0})(z_{\mathrm{r}}-L)} = \frac{\ell_{\mathrm{sca}}(\omega_0)}{z_{\mathrm{r}}-L}\omega_0^2,$$

and since we consider situations in which the depth $z_{\mathrm{r}} - L$ of the reflector is of the same order as $\ell_{\mathrm{sca}}(\omega_0)$, we can see that the frequency coherence radius is of the order of the central frequency.

5

Detection, localization, and characterization of reflectors by random matrix theory

In sensor array imaging, waves are emitted by a set of sources and they are recorded by a set of sensors in order to probe an unknown medium. The responses between each pair of source and receiver are collected and assembled in the form of a matrix, the array response matrix. The indices of the matrix are the index of the source and the index of the receiver. When the data are corrupted by additive noise, we study the structure of the array response matrix using random matrix theory. We start this chapter by presenting an acquisition scheme, known as Hadamard technique, for noise reduction. Hadamard technique allows us to acquire simultaneously the elements of the array response matrix. The feature of this technique is to divide the variance of the noise by the number of sources. Then we investigate the statistical distributions of the singular values of the array response matrix in the presence of point reflectors. These results can be used to design and calibrate detection tests, to estimate the number of point reflectors point reflector in the medium, to localize them, and to estimate their reflectivities.

5.1 Data acquisition and Hadamard technique

We consider a situation similar to the one introduced in Section 3.1. We actually introduce a generalized set up in which the sources and receivers are not coincident and there may be several reflectors. We consider point reflector the Helmholtz equation:

$$\Delta_{\boldsymbol{z}}\hat{u}(\omega,\boldsymbol{z},\boldsymbol{x})+\frac{\omega^2}{c_o^2}\Big(1+\sum_{j=1}^{r}V_j(\boldsymbol{z})\Big)\hat{u}(\omega,\boldsymbol{z},\boldsymbol{x})=-\delta(\boldsymbol{z}-\boldsymbol{x})\quad\text{in }\mathbb{R}^3 \tag{5.1}$$

for $\boldsymbol{x}\in\mathbb{R}^3$, with the Sommerfeld radiation condition imposed on $\hat{u}$. Here r is the number of localized reflectors, $\boldsymbol{x}$ is the location of the source, and

$$V_j(\boldsymbol{z})=\eta_j\mathbf{1}_{\Omega_j}(\boldsymbol{z}-\boldsymbol{z}_j)\;, \tag{5.2}$$

where, for $j = 1, \ldots, r$, Ω_j is a compactly supported domain with volume $|\Omega_j|$ that describes the shape of the jth inclusion, $\boldsymbol{z}_j$ is for the location of the jth inclusion, and η_j is the contrast of the jth inclusion.

Suppose that we have a transmitter array of N_s sources located at $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{N_\mathrm{s}}\}$ and a receiver array of N_r elements located at $\{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_{N_\mathrm{r}}\}$. The $N_\mathrm{r} \times N_\mathrm{s}$ response matrix $\mathbf{A}$ describes the transmit-receive process performed at these arrays. The field received by the nth receiving element $\boldsymbol{y}_n$ when a unit-amplitude time-harmonic wave is emitted from $\boldsymbol{x}_m$ is $\hat{u}(\omega, \boldsymbol{y}_n, \boldsymbol{x}_m)$. If we remove the incident field then we obtain the (n, m)-th entry of the unperturbed, perfect array response matrix $\mathbf{A}^0$:

$$A^0_{nm} = \hat{u}(\omega, \boldsymbol{y}_n, \boldsymbol{x}_m) - \hat{G}_0(\omega, \boldsymbol{y}_n, \boldsymbol{x}_m) \,. \tag{5.3}$$

The incident field is $\hat{G}_0(\omega, \boldsymbol{z}, \boldsymbol{x}_m)$. However noise cannot be neglected in the acquisition process, so we actually record a perturbed version of $\mathbf{A}^0$ whose form depends on the acquisition scheme.

In the standard acquisition scheme, the response matrix is measured during a sequence of N_s experiments. In the mth experiment, $m = 1, \ldots, N_\mathrm{s}$, the mth source generates the incident field (i.e. only the mth source emits, and it emits a unit-amplitude time-harmonic signal) and the N_r receivers record the scattered wave which means that they measure

$$A^{\mathrm{meas}}_{nm} = A^0_{nm} + W_{nm}, \quad n = 1, \ldots, N_\mathrm{r}, \quad m = 1, \ldots, N_\mathrm{s} \,,$$

which gives the matrix

$$\mathbf{A}^{\mathrm{meas}} = \mathbf{A}^0 + \mathbf{W} \,, \tag{5.4}$$

where $\mathbf{A}^0$ is the unperturbed response matrix (5.3) and W_{nm} are independent complex Gaussian random variables with mean zero and variance $\sigma^2_{\mathrm{noise}}$.

The Hadamard technique is a noise reduction technique in the presence of additive noise that uses the structure of Hadamard matrices. It allows to acquire the elements of the array response matrix simultaneously.

Definition 5.1. *A real Hadamard matrix* $\mathbf{H}$ *of order* N_s *is an* $N_\mathrm{s} \times N_\mathrm{s}$ *matrix whose elements are* -1 *or* $+1$ *and such that* $\mathbf{H}^T\mathbf{H} = N_\mathrm{s}\mathbf{I}$.

Real Hadamard matrices do not exist for all N_s. A necessary condition for the existence is that $N_\mathrm{s} = 1, 2$ or a multiple of 4. A sufficient condition is that N_s is a power of two. Explicit examples are known for all N_s multiple of 4 up to $N_\mathrm{s} = 664$ [36].

Definition 5.2. *A complex Hadamard matrix* $\mathbf{H}$ *of order* N_s *is an* $N_\mathrm{s} \times N_\mathrm{s}$ *matrix whose elements have modulus one and such that* $\mathbf{H}^\dagger\mathbf{H} = N_\mathrm{s}\mathbf{I}$.

Here † denotes the conjugate transpose. Complex Hadamard matrices exist for all N_s. An example is the Fourier matrix $H_{nm} = \exp(i2\pi(n-1)(m-1)/N_\mathrm{s})$. The important result is the following one: A Hadamard matrix has maximal

determinant among matrices with complex entries in the closed unit disk. More exactly the determinant of any complex $N_s \times N_s$ matrix $\mathbf{H}$ with $|H_{nm}| \leq 1$ for all $n, m = 1, \ldots, N_s$ satisfies $|\det \mathbf{H}| \leq N_s^{N_s/2}$, with equality attained by a complex Hadamard matrix. This result has important consequences as we describe below.

Hadamard acquisition. We now describe a general multi-source acquisition scheme and show the importance of Hadamard matrices to build an optimal scheme. Let $\mathbf{H}$ be a complex inversible $N_s \times N_s$ matrix with $|H_{nm}| \leq 1$ for all $n, m = 1, \ldots, N_s$. In the multi-source acquisition scheme, the response matrix is measured during a sequence of N_s experiments:

- In the mth experience, $m = 1, \ldots, N_s$, all sources generate time-harmonic signals, the m'th source generating $H_{m'm}$ (the amplitude is bounded by one). This means that we use all sources to their maximal emission capacity with a specific coding of their phases.
- The N_r receivers record the scattered wave which means that the nth receiver records

$$B^{\text{meas}}_{nm} = \sum_{m'=1}^{N_s} A^0_{nm'} H_{m'm} + W_{nm} = (\mathbf{A}^0\mathbf{H})_{nm} + W_{nm}$$

for $n = 1, \ldots, N_r$, and $m = 1, \ldots, N_s$, which gives the matrix

$$\mathbf{B}^{\text{meas}} = \mathbf{A}^0\mathbf{H} + \mathbf{W} \,,$$

where $\mathbf{A}^0$ is the unperturbed response matrix (5.3) and W_{nm} are independent Gaussian random variables with mean zero and variance σ^2_{noise}. The measured response matrix $\mathbf{A}^{\text{meas}}$ is obtained by right multiplying the matrix $\mathbf{B}^{\text{meas}}$ by the matrix $\mathbf{H}^{-1}$:

$$\mathbf{A}^{\text{meas}} := \mathbf{B}^{\text{meas}}\mathbf{H}^{-1} = \mathbf{A}^0 + \widetilde{\mathbf{W}}, \qquad \widetilde{\mathbf{W}} = \mathbf{W}\mathbf{H}^{-1} \,. \tag{5.5}$$

The choice of the matrix $\mathbf{H}$ should fulfill the property that the new noise matrix $\widetilde{\mathbf{W}} = \mathbf{W}\mathbf{H}^{-1}$ has independent complex entries with mean zero and minimal variance. We have

$$\begin{aligned}
\mathbb{E}\Big[\overline{\widetilde{W}_{nm}}\widetilde{W}_{n'm'}\Big] &= \sum_{q,q'=1}^{N_s} \overline{(\mathbf{H}^{-1})_{qm}}(\mathbf{H}^{-1})_{q'm'}\mathbb{E}\big[\overline{W_{nq}}W_{n'q'}\big] \\
&= \sigma^2_{\text{noise}} \sum_{q,q'=1}^{N_s} \overline{(\mathbf{H}^{-1})_{qm}}(\mathbf{H}^{-1})_{q'm'}\mathbf{1}_n(n')\mathbf{1}_q(q') \\
&= \sigma^2_{\text{noise}} \sum_{q=1}^{N_s} ((\mathbf{H}^{-1})^\dagger)_{mq}(\mathbf{H}^{-1})_{qm'}\mathbf{1}_n(n') \\
&= \sigma^2_{\text{noise}}((\mathbf{H}^{-1})^\dagger\mathbf{H}^{-1})_{mm'}\mathbf{1}_n(n').
\end{aligned}$$

This shows that we look for a complex matrix $\mathbf{H}$ with entries in the unit disk such that $(\mathbf{H}^{-1})^\dagger\mathbf{H}^{-1} = c\mathbf{I}$ with a minimal c. This is equivalent to require

that $\mathbf{H}$ is proportional to a unitary matrix and that $|\det \mathbf{H}|$ is maximal. Using Hadamard result we know that the maximal determinant is $N_{\mathrm{s}}^{N_{\mathrm{s}}/2}$ and that a complex Hadamard matrix attains the maximum. In other words, the optimal matrix $\mathbf{H}$ that minimizes the noise variance should be a Hadamard matrix. The feature of the Hadamard technique is that the new noise matrix $\widetilde{\mathbf{W}}$ has independent entries with Gaussian statistics, mean zero, and variance $\sigma^2_{\mathrm{noise}}/N_{\mathrm{s}}$:

$$\mathbb{E}\big[\widetilde{W}_{nm}\widetilde{W}_{n'm'}\big] = \frac{\sigma^2_{\mathrm{noise}}}{N_{\mathrm{s}}}\delta_{mm'}\delta_{nn'} \ .$$

This gain of a factor N_{s} in the signal-to-noise ratio (compared to (5.4)) is called the Hadamard advantage.

5.2 Singular Value Decomposition of Array Response Matrices

5.2.1 Point Reflectors

Taking into account measurement noise, the measured response matrix $\mathbf{A}^{\mathrm{meas}}$ is

$$\mathbf{A}^{\mathrm{meas}} = \mathbf{A}^0 + \frac{1}{\sqrt{N_{\mathrm{s}}}}\mathbf{W} \ , \tag{5.6}$$

where $\mathbf{A}^0$ is the unperturbed response matrix (5.3) and the matrix $\mathbf{W}$ represents the additive measurement noise, which is a random matrix with independent and identically distributed complex entries with Gaussian statistics, mean zero and variance $\sigma^2_{\mathrm{noise}}$. This particular scaling for the noise level is the right one to get non-trivial asymptotic regimes in the limit $N_{\mathrm{s}} \to \infty$. Furthermore, it is the regime that emerges from the use of the Hadamard acquisition scheme for the array response matrix as described above.

In the Born approximation, where the volume $|\Omega_j|$ of $\Omega_j, j = 1, \ldots, r$, is small, the measured field has approximately the following form:

$$\hat{u}(\omega, \boldsymbol{y}_n, \boldsymbol{x}_n) \approx \hat{G}_0(\omega, \boldsymbol{y}_n, \boldsymbol{x}_n) + \sum_{j=1}^{r} \rho_j \hat{G}_0(\omega, \boldsymbol{y}_n, \boldsymbol{z}_j)\hat{G}_0(\omega, \boldsymbol{z}_j, \boldsymbol{x}_n), \tag{5.7}$$

for $n = 1, \ldots, N_{\mathrm{r}}$, $m = 1, \ldots, N_{\mathrm{s}}$, where ρ_j is the coefficient of reflection defined by

$$\rho_j = \frac{\omega^2}{c_o^2}\eta_j|\Omega_j| \ . \tag{5.8}$$

We introduce the normalized vector of fundamental solutions from the receiver array to the point $\boldsymbol{z}$:

$$\boldsymbol{w}(\boldsymbol{z}) := \frac{1}{\big(\sum_{l=1}^{N_{\mathrm{r}}} |\hat{G}_0(\omega, \boldsymbol{z}, \boldsymbol{y}_l)|^2\big)^{\frac{1}{2}}} \big(\hat{G}_0(\omega, \boldsymbol{z}, \boldsymbol{y}_n)\big)_{n=1,\ldots,N_{\mathrm{r}}} \ , \tag{5.9}$$

and the normalized vector of fundamental solutions from the transmitter array to the point $\boldsymbol{z}$, known as the illumination vector,

$$\boldsymbol{v}(\boldsymbol{z}) := \frac{1}{\big(\sum_{l=1}^{N_\mathrm{s}} |\hat{G}_0(\omega, \boldsymbol{z}, \boldsymbol{x}_l)|^2\big)^{\frac{1}{2}}} \big(\overline{\hat{G}_0(\omega, \boldsymbol{z}, \boldsymbol{x}_m)}\big)_{m=1,\ldots,N_\mathrm{s}} . \tag{5.10}$$

Using (5.7) we can then write the unperturbed response matrix in the form

$$\mathbf{A}^0 = \sum_{j=1}^{r} \sigma_j \boldsymbol{w}(\boldsymbol{z}_j)\boldsymbol{v}(\boldsymbol{z}_j)^\dagger , \tag{5.11}$$

with

$$\sigma_j := \rho_j \Big(\sum_{n=1}^{N_\mathrm{r}} |\hat{G}_0(\omega, \boldsymbol{z}_j, \boldsymbol{y}_n)|^2\Big)^{\frac{1}{2}} \Big(\sum_{m=1}^{N_\mathrm{s}} |\hat{G}_0(\omega, \boldsymbol{z}_j, \boldsymbol{x}_m)|^2\Big)^{\frac{1}{2}} . \tag{5.12}$$

Let us assume for a while that the arrays of transmitters and receivers are equi-distributed on a sphere englobing the point reflectors. Moreover, the point reflectors are at a distance from the arrays of transmitters and receivers much larger than the wavelength $2\pi c_o/\omega$. Provided that the positions $\boldsymbol{z}_j$ of the reflectors are far from one another or well-separated (*i.e.*, farther than the wavelength $2\pi c_o/\omega$), the vectors $\boldsymbol{w}(\boldsymbol{z}_j)$, $j = 1, \ldots, r$, are approximately orthogonal to one another, as well as are the vectors $\boldsymbol{v}(\boldsymbol{z}_j)$, $j = 1, \ldots, r$. In fact, from the Helmholtz-Kirchhoff identity (1.11), we have

$$\frac{1}{N_\mathrm{r}} \sum_{n=1}^{N_\mathrm{r}} \hat{G}_0(\omega, \boldsymbol{z}_j, \boldsymbol{y}_n)\overline{\hat{G}_0(\omega, \boldsymbol{z}_i, \boldsymbol{y}_n)} \approx \mathrm{sinc}\Big(\frac{\omega|\boldsymbol{z}_i - \boldsymbol{z}_j|}{c_o}\Big) \tag{5.13}$$

as $N_\mathrm{r} \to +\infty$. Moreover, $\mathrm{sinc}(\omega|\boldsymbol{z}_i - \boldsymbol{z}_j|/c_o) \approx 0$ when $|\boldsymbol{z}_j - \boldsymbol{z}_i|$ is much larger than the wavelength. The matrix $\mathbf{A}^0$ then has rank r and its nonzero singular values are σ_j, $j = 1, \ldots, r$, with the associated left and right singular vectors $\boldsymbol{w}(\boldsymbol{z}_j)$ and $\boldsymbol{v}(\boldsymbol{z}_j)$.

The previous observation can be extended beyond the situation where the arrays of transmitters and receivers are equi-distributed on a sphere, and it is a general result that the matrix $\mathbf{A}^0$ has rank r, singular values are σ_j, and singular vectors $\boldsymbol{w}(\boldsymbol{z}_j)$ and $\boldsymbol{v}(\boldsymbol{z}_j)$, as soon as the reflectors are well-separated. In these conditions the following proposition turns out to be very useful.

Proposition 5.3. *Let $\mathbf{A}^0$ be a $N_\mathrm{r} \times N_\mathrm{s}$ matrix with rank r. Let us denote $\sigma_1(\mathbf{A}^0) \geq \cdots \geq \sigma_r(\mathbf{A}^0) > 0$ its nonzero singular values. Let $\mathbf{W}$ be a $N_\mathrm{r} \times N_\mathrm{s}$ random matrix with independent and identically distributed complex entries with Gaussian statistics, mean zero, and variance $\sigma^2_{\mathrm{noise}}$. We define $\mathbf{A}^{\mathrm{meas}}$ by (5.6). When $\gamma = N_\mathrm{r}/N_\mathrm{s}$ is fixed and $N_\mathrm{s} \to \infty$, the singular values of the random matrix $\mathbf{A}^{\mathrm{meas}}$ satisfy*

$$\sigma_j(\mathbf{A}^{\mathrm{meas}}) \overset{N_\mathrm{s}\to\infty}{\longrightarrow} \begin{cases} \sigma_{\mathrm{noise}}\Big(\dfrac{\sigma_j^2(\mathbf{A}^0)}{\sigma_{\mathrm{noise}}^2} + 1 + \gamma + \gamma\dfrac{\sigma_{\mathrm{noise}}^2}{\sigma_j^2(\mathbf{A}^0)}\Big)^{\frac{1}{2}} & \text{if } \sigma_j(\mathbf{A}^0) > \gamma^{\frac{1}{4}}\sigma_{\mathrm{noise}}\ , \\ \sigma_{\mathrm{noise}}(1+\gamma^{\frac{1}{2}}) & \text{if } \sigma_j(\mathbf{A}^0) \le \gamma^{\frac{1}{4}}\sigma_{\mathrm{noise}} \end{cases} \tag{5.14}$$

for any $j = 1, \dots, r$, in probability.

Proposition 5.3 shows how the singular values of the perturbed response matrix $\mathbf{A}^{\mathrm{meas}}$ are related to the singular values of the unperturbed response matrix $\mathbf{A}^0$. We can see that there is level repulsion for the singular values $\sigma_j(\mathbf{A}^0)$ that are larger than the threshold value $\gamma^{1/4}\sigma_{\mathrm{noise}}$, in the sense that $\sigma_j(\mathbf{A}^{\mathrm{meas}}) > \sigma_j(\mathbf{A}^0)$. We can also observe that the singular values $\sigma_j(\mathbf{A}^0)$ that are smaller than the threshold value $\gamma^{1/4}\sigma_{\mathrm{noise}}$ are absorbed in the deformed quarter-circle distribution of the singular values of the noise matrix $\mathbf{W}/\sqrt{N_\mathrm{s}}$, which extends up to $\sigma_{\mathrm{noise}}(1+\gamma^{1/2})$ (see Proposition 5.6).

5.2.2 SVD of a Gaussian Random Matrix

As a warm-up we consider first an $N_\mathrm{r} \times N_\mathrm{s}$ real matrix $\mathbf{A}$ with $N_\mathrm{r} \ge N_\mathrm{s}$. We assume that $\mathbf{A}$ consists of independent Gaussian noise entries with mean zero and variance $\sigma_{\mathrm{noise}}^2/N_\mathrm{s}$. We denote by $\sigma_1^{(N_\mathrm{s})} \ge \sigma_2^{(N_\mathrm{s})} \ge \sigma_3^{(N_\mathrm{s})} \ge \cdots \ge \sigma_{N_\mathrm{s}}^{(N_\mathrm{s})}$ the singular values of the matrix $\mathbf{A}$ sorted in decreasing order and by $\Lambda^{(N_\mathrm{s})}$ the corresponding density of states defined by

$$\Lambda^{(N_\mathrm{s})}([a,b]) := \frac{1}{N_\mathrm{s}}\mathrm{Card}\left\{l = 1, \dots, N_\mathrm{s}\ ,\ \sigma_l^{(N_\mathrm{s})} \in [a,b]\right\} \quad \text{for any } a < b\ .$$

The density of states $\Lambda^{(N_\mathrm{s})}$ is a counting measure which consists of a sum of Dirac masses:

$$\Lambda^{(N_\mathrm{s})} = \frac{1}{N_\mathrm{s}}\sum_{j=1}^{N_\mathrm{s}} \delta_{\sigma_j^{(N_\mathrm{s})}}\ .$$

For large N_r and N_s with $N_\mathrm{r}/N_\mathrm{s} = \gamma \ge 1$ fixed we have the following results.

Proposition 5.4. *Let $\mathbf{A}$ be an $N_\mathrm{r} \times N_\mathrm{s}$ Gaussian real matrix.*

(i) *The random measure $\Lambda^{(N_\mathrm{s})}$ almost surely converges to the deterministic absolutely continuous measure Λ_γ with compact support:*

$$\Lambda_\gamma([\sigma_u, \sigma_v]) = \int_{\sigma_u}^{\sigma_v} \frac{1}{\sigma_{\mathrm{noise}}}\rho_\gamma\Big(\frac{\sigma}{\sigma_{\mathrm{noise}}}\Big)d\sigma, \qquad 0 \le \sigma_u \le \sigma_v\ , \tag{5.15}$$

where ρ_γ is the deformed quarter-circle law given by

$$\rho_\gamma(\sigma) = \begin{cases} \dfrac{1}{\pi\sigma}\sqrt{\big((\gamma^{1/2}+1)^2 - \sigma^2\big)\big(\sigma^2 - (\gamma^{1/2}-1)^2\big)} \\ \qquad \text{if } \gamma^{1/2} - 1 \le \sigma \le \gamma^{1/2} + 1\ , \\ 0 \qquad \text{otherwise.} \end{cases} \tag{5.16}$$

(ii) *The normalized l^2-norm of the singular values satisfies*

$$N_{\mathrm{s}}\Big[\frac{1}{N_{\mathrm{s}}}\sum_{j=1}^{N_{\mathrm{s}}}(\sigma_j^{(N_{\mathrm{s}})})^2-\gamma\sigma^2_{\mathrm{noise}}\Big]\overset{N_{\mathrm{s}}\to\infty}{\longrightarrow}\sqrt{2\gamma}\sigma^2_{\mathrm{noise}}Z_0 \text{ in distribution}, \quad (5.17)$$

where Z_0 follows a Gaussian distribution with mean zero and variance one.

(iii) *The maximal singular value satisfies*

$$\sigma_1^{(N_{\mathrm{s}})}\approx\sigma_{\mathrm{noise}}\Big[\gamma^{1/2}+1+\frac{1}{2N_{\mathrm{s}}^{2/3}}\big(1+\gamma^{-1/2}\big)^{1/3}Z_1+o(\frac{1}{N_{\mathrm{s}}^{2/3}})\Big] \text{ in distribution}, \quad (5.18)$$

where Z_1 follows a type-1 Tracy-Widom distribution.

Proof. The type-1 Tracy-Widom distribution has the pdf p_{TW1} defined in Appendix 6.6. (i) is the so-called Marcenko-Pastur result [33]. (ii) follows from the expression of the normalized l^2-norm of the singular values in terms of the entries of the matrix:

$$\frac{1}{N_{\mathrm{s}}}\sum_{j=1}^{N_{\mathrm{s}}}(\sigma_j^{(N_{\mathrm{s}})})^2=\frac{1}{N_{\mathrm{s}}}\mathrm{Tr}\big(\mathbf{A}^T\mathbf{A}\big)=\frac{1}{N_{\mathrm{s}}}\sum_{j=1}^{N_{\mathrm{r}}}\sum_{l=1}^{N_{\mathrm{s}}}A_{jl}^2\,,$$

and from the application of the central limit theorem in the regime $N_{\mathrm{s}}\gg 1$. (iii) follows from [29]. □

A similar result can be obtained when the matrix $\mathbf{A}$ is an $N_{\mathrm{r}}\times N_{\mathrm{s}}$ complex-valued matrix with $N_{\mathrm{r}}\geq N_{\mathrm{s}}$ and the coefficients of $\mathbf{A}$ are independent complex Gaussian random variables with mean zero and variance $\sigma^2_{\mathrm{noise}}/N_{\mathrm{s}}$ (which means that the real and imaginary parts are independent and identically distributed Gaussian random variables with mean zero and variance $\sigma^2_{\mathrm{noise}}/(2N_{\mathrm{s}})$). Then the previous proposition holds, except that the limit of the normalized l^2-norm is $\sqrt{\gamma}\sigma^2_{\mathrm{noise}}Z_0$ (instead of $\sqrt{2\gamma}\sigma^2_{\mathrm{noise}}Z_0$) and the fluctuations of the maximal singular value are described in terms of a type-2 Tracy-Widom distribution (instead of a type-1 Tracy-Widom distribution) described in Appendix 6.6.

Proposition 5.5. *Let $\mathbf{A}$ be an $N_{\mathrm{r}}\times N_{\mathrm{s}}$ Gaussian complex matrix.*

(i) *The random measure $\Lambda^{(N_{\mathrm{s}})}$ almost surely converges to the deterministic absolutely continuous measure Λ_γ defined by (5.15).*

(ii) *The normalized l^2-norm of the singular values satisfies*

$$N_{\mathrm{s}}\Big[\frac{1}{N_{\mathrm{s}}}\sum_{j=1}^{N_{\mathrm{s}}}(\sigma_j^{(N_{\mathrm{s}})})^2-\gamma\sigma^2_{\mathrm{noise}}\Big]\overset{N_{\mathrm{s}}\to\infty}{\longrightarrow}\sqrt{\gamma}\sigma^2_{\mathrm{noise}}Z_0 \text{ in distribution}, \quad (5.19)$$

where Z_0 follows a Gaussian distribution with mean zero and variance one.

(iii) *The maximal singular value satisfies*

$$\sigma_1^{(N_\mathrm{s})} = \sigma_\mathrm{noise}\Big[\gamma^{1/2}+1+\frac{1}{2N_\mathrm{s}^{2/3}}\big(1+\gamma^{-1/2}\big)^{1/3}Z_2+o(\frac{1}{N_\mathrm{s}^{2/3}})\Big] \text{ in distribution,} \tag{5.20}$$

where Z_2 follows a type-2 Tracy-Widom distribution.

This proposition describes the distribution of the singular values of the array response matrix in the absence of reflector.

The response matrix using the Hadamard technique in the presence of a point reflector and in the presence of measurement noise is

$$\mathbf{A} = \mathbf{A}^0 + \frac{1}{\sqrt{N_\mathrm{s}}}\mathbf{W}, \tag{5.21}$$

where $\mathbf{A}^0$ is a rank-one matrix with a unique non-zero singular value σ_0 and $\mathbf{W}$ has independent random complex entries with Gaussian statistics, mean zero and variance σ^2_noise. We still denote by $\sigma_1^{(N_\mathrm{s})} \geq \sigma_2^{(N_\mathrm{s})} \geq \sigma_3^{(N_\mathrm{s})} \geq \cdots \geq \sigma_{N_\mathrm{s}}^{(N_\mathrm{s})}$ the singular values of the matrix $\mathbf{A}$ sorted in decreasing order.

Proposition 5.6. *In the regime $N_\mathrm{s} \to \infty$:*

a) *The normalized l^2-norm of the singular values satisfies*

$$N_\mathrm{s}\Big[\frac{1}{N_\mathrm{s}}\sum_{j=1}^{N_\mathrm{s}}(\sigma_j^{(N_\mathrm{s})})^2 - \gamma\sigma^2_\mathrm{noise}\Big] \overset{N_\mathrm{s}\to\infty}{\longrightarrow} \sigma_0^2 + \sqrt{2\gamma}\sigma^2_\mathrm{noise}Z_0 \text{ in distribution,} \tag{5.22}$$

where Z_0 follows a Gaussian distribution with mean zero and variance one.

b1) *If $\sigma_0 < \gamma^{1/4}\sigma_\mathrm{noise}$, then the maximal singular value satisfies*

$$\sigma_1^{(N_\mathrm{s})} = \sigma_\mathrm{noise}\Big[\gamma^{1/2}+1+\frac{1}{2N_\mathrm{s}^{2/3}}\big(1+\gamma^{-1/2}\big)^{1/3}Z_2+o(\frac{1}{N_\mathrm{s}^{2/3}})\Big] \text{ in distribution,} \tag{5.23}$$

where Z_2 follows a type-2 Tracy-Widom distribution.

b2) *If $\sigma_0 > \gamma^{1/4}\sigma_\mathrm{noise}$, then the maximal singular value has Gaussian distribution with the mean and variance given by*

$$\mathbb{E}\big[\sigma_1^{(N_\mathrm{s})}\big] = \sigma_0\Big[\Big(1+(1+\gamma)\frac{\sigma^2_\mathrm{noise}}{\sigma_0^2}+\gamma\frac{\sigma^4_\mathrm{noise}}{\sigma_0^4}\Big)^{1/2}+o(\frac{1}{N_\mathrm{s}^{1/2}})\Big], \tag{5.24}$$

$$\mathrm{Var}\big(\sigma_1^{(N_\mathrm{s})}\big) = \frac{\sigma^2_\mathrm{noise}}{2N_\mathrm{s}}\Big[\frac{1-\gamma\frac{\sigma^4_\mathrm{noise}}{\sigma_0^4}}{\big(1+(1+\gamma)\frac{\sigma^2_\mathrm{noise}}{\sigma_0^2}+\gamma\frac{\sigma^4_\mathrm{noise}}{\sigma_0^4}\big)^{1/2}}+o(1)\Big]. \tag{5.25}$$

This proposition describes the distribution of the singular values of the array response matrix in the presence of a reflector.

5.3 Detection Test

The objective of this subsection is to design specific point reflector detection rules. To simplify the presentation, we assume from now on that the sources and receivers are coincident so that $N_{\mathrm{r}} = N_{\mathrm{s}} = N$.

5.3.1 SVD Based Detection Test

Suppose that the array response matrix $\mathbf{A}^{\mathrm{meas}}$ consists of independent Gaussian noise coefficients with mean zero and variance $\sigma^2_{\mathrm{noise}}/N$. Let the ratio R of the first singular value $\sigma_1(\mathbf{A}^{\mathrm{meas}})$ over the normalized l^2-norm of the other singular values $(\sigma_j(\mathbf{A}^{\mathrm{meas}}))_{j=2,\dots,N}$ of the measured array response matrix $\mathbf{A}^{\mathrm{meas}}$ be defined by

$$R := \frac{\sigma_1(\mathbf{A}^{\mathrm{meas}})}{\left(\frac{1}{N-1}\sum_{j=2}^{N}\sigma_j(\mathbf{A}^{\mathrm{meas}})^2\right)^{1/2}} \,. \tag{5.26}$$

Using Proposition 5.6 (b and c) and Slutsky's theorem, we obtain the following result.

Proposition 5.7. *In the absence of any point reflector, the ratio R defined by (5.26) has the following statistical distribution*

$$R \overset{dist.}{=} 2 + \frac{1}{2^{2/3}N^{2/3}} Z_2 \,, \tag{5.27}$$

when N is large, where Z_2 is a random variable following a type 2 Tracy-Widom distribution.

This proposition describes the statistical distribution of the ratio (5.26) in the absence of a point reflector. As we will see, it allows us to compute explicitly the threshold of a likelihood-ratio test.

Now we turn to the case where the array response matrix is obtained with a single point reflector in the presence of additive noise. Then,

$$\mathbf{A}^{\mathrm{meas}} = \mathbf{A}^0 + \frac{1}{\sqrt{N}}\mathbf{W} \,,$$

where $\mathbf{A}^0$ is the unperturbed array response matrix (5.11) with $r = 1$ (corresponding to one point reflector) and the entries of the matrix $\mathbf{W}$ are independent complex Gaussian random variables with mean zero and variance $\sigma^2_{\mathrm{noise}}$. Let σ_0 be the nonzero singular value of $\mathbf{A}^0$ and let $\sigma_1(\mathbf{A}^{\mathrm{meas}}) \geq \sigma_2(\mathbf{A}^{\mathrm{meas}}) \geq \cdots \geq \sigma_N(\mathbf{A}^{\mathrm{meas}})$ be the singular values of the measured array response matrix $\mathbf{A}^{\mathrm{meas}}$. Using Propositions 5.3 and 5.6 we can describe the ratio of the maximal singular value over the normalized l^2-norm as follows.

Proposition 5.8. *Let us consider the array response matrix obtained in the presence of measurement noise with a point reflector. For* $\sigma_{\text{noise}} < \sigma_0$*, the ratio* R *defined by (5.26) has the following statistical distribution*

$$R \overset{dist.}{=} \frac{\sigma_0}{\sigma_{\text{noise}}} + \frac{\sigma_{\text{noise}}}{\sigma_0} + \frac{1}{\sqrt{2N}} \sqrt{1 - \sigma_{\text{noise}}^2 \sigma_0^{-2}} Z_0 \,, \tag{5.28}$$

where Z_0 *follows a Gaussian distribution with mean zero and variance one. For* $\sigma_{\text{noise}} > \sigma_0$ *we have (5.27).*

This proposition describes the statistical distribution of the ratio (5.26) in the presence of a point reflector. It allows us to compute explicitly the power of the likelihood-ratio test which is the most powerful test for a given false alarm rate by the Neyman-Pearson lemma.

5.3.2 Statistical Test

As in the standard statistical hypothesis testing [18, 30], we postulate two hypotheses and derive a decision rule for deciding in between them based on the measured array response matrix.

We define H_o the (null) hypothesis to be tested and H_a the (alternative) hypothesis:

- H_o: there is no point reflector,
- H_a: there is a point reflector.

We want to test H_o against H_a. Two types of independent errors can be made:

- Type I errors correspond to rejecting the null hypothesis H_o when it is correct (false alarm). Their probability is given by

$$\alpha := \mathbb{P}\big(\text{accept } H_a | H_o \text{ true}\big) \,.$$

- Type II errors correspond to accepting H_o when it is false (missed detection) and have probability

$$\beta := \mathbb{P}\big(\text{accept } H_o | H_a \text{ true}\big) \,.$$

The success of the test (probability of detection, POD, or detection power) is therefore given by $1 - \beta$.

Given the data the decision rule for accepting H_o or not can be derived from the Neyman-Pearson lemma which asserts that for a prescribed false alarm rate α the most powerful test corresponds to accepting H_a for the likelihood ratio of H_a to H_o exceeding a threshold value determined by α.

Neyman-Pearson Lemma: Let $\mathcal{Y}$ be the set of all possible data and let $f_0(y)$ and $f_1(y)$ be the probability densities of Y under the null and alternative hypotheses. The Neyman-Pearson lemma [18, p. 335] states that the most powerful test has a critical region defined by

$$\mathcal{Y}_\alpha := \left\{ y \in \mathcal{Y} \;\middle|\; \frac{f_1(y)}{f_0(y)} \geq \eta_\alpha \right\} , \tag{5.29}$$

for a threshold η_α satisfying

$$\int_{\mathcal{Y}_\alpha} f_0(y)dy = \alpha . \tag{5.30}$$

Let the data be y. We reject H_o if the likelihood ratio $\frac{f_1(y)}{f_0(y)} > \eta_\alpha$ and accept H_o otherwise. The power of the (most powerful) test is

$$1 - \beta = \int_{\mathcal{Y}_\alpha} f_1(y)dy . \tag{5.31}$$

5.3.3 Berens' Modeling

In [6] a framework for analyzing schemes for nondestructive inspection methods and testing for the presence of flaws was introduced. In this reliability analysis the probability of detection (POD) as a function of flaw size played a central role. In our notation the "flaw size" corresponds to the parameter ρ (given by (5.8)) and we are thus interested in designing reliability tests with a desirable performance in terms of the corresponding POD(ρ) function. In [6] a maximum likelihood approach was used for parameter estimation, and a log normal distribution was in particular postulated for the response variable's relation to point reflector strength. One parameter to be estimated is then the variance of the Gaussian residual. Our approach here is to introduce a physical model for the measurements, as we have described above, and then infer a corresponding "optimal" POD function that can be associated with the array response matrix measurements. We describe the picture deriving from this approach below. It turns out that the resulting picture deviates somewhat from the one derived from Berens' modeling.

Consider the imaging of point reflectors from measurements of the array response matrix at a single frequency ω in the presence of measurement noise, that is, we model with an additive Gaussian noise. Assuming availability of previous and/or multiple measurements we may assume that the variance of the entries of the array response matrix (due to the measurement noise) is known and equal to σ^2_{noise}. In fact, we will see that we do not need to know the value σ^2_{noise} in order to build the most powerful test with a prescribed false alarm rate.

In the absence of the point reflector (hypothesis H_o) the statistical distribution of the ratio R of the first singular value of the symmetrized array response matrix over the normalized l^2-norm of the other singular values is of the form (5.27).

In the presence of a point reflector at position $\boldsymbol{z}_0$ and with coefficient of reflection ρ_0 (hypothesis H_a), Proposition 5.8 shows that the ratio is of the form (5.28), with σ_0 given by (5.12):

$$\sigma_0 = \rho_0 \Big(\sum_{n=1}^{N} |\hat{G}_0(\omega, \boldsymbol{z}_0, \boldsymbol{x}_n)|^2 \Big) \ . \tag{5.32}$$

This result is correct as long as $\sigma_0 > \sigma_{\text{noise}}$. When $\sigma_0 < \sigma_{\text{noise}}$ we have (5.27).

If the data gives the ratio R, then we propose to use a test of the form $R > r$ for the alarm corresponding to the presence of a point reflector. By the Neyman-Pearson lemma the decision rule of accepting H_a if and only if $R > r_\alpha$ maximizes the probability of detection for a given false alarm probability α

$$\alpha = \mathbb{P}(R > r_\alpha | H_o) \ ,$$

with the threshold

$$r_\alpha = 2 + \frac{1}{2^{2/3} N^{2/3}} \Phi_{\text{TW2}}^{-1}(1-\alpha) \ , \tag{5.33}$$

where $\Phi_{\text{TW2}}(x) = \int_{-\infty}^{x} p_{\text{TW2}}(y) dy$ is the cumulative distribution function of the Tracy-Widom distribution of type 2. The computation of the threshold is easy since it depends only on the number of sensors N and on the false alarm probability α. This test is therefore universal. Note that we should use a Tracy-Widom distribution table, and not a Gaussian table. We have, for instance, $\Phi_{\text{TW2}}^{-1}(0.99) \approx 0.48$.

The detection probability $1-\beta$ is the probability to sound the alarm when there is a point reflector:

$$1 - \beta = \mathbb{P}(R > r_\alpha | H_a) \ .$$

For a given measurement array it depends on ρ_0 and $\boldsymbol{z}_0$ through the value σ_0 defined by (5.32) and also on the noise level σ_{noise}. Here we find that the Probability Of Detection is

$$\text{POD} = \Phi \left(\frac{\sqrt{N} \frac{\sigma_0}{\sigma_{\text{noise}}} + \frac{\sigma_{\text{noise}}}{\sigma_0} - r_\alpha}{\sqrt{1 - (\sigma_{\text{noise}}/\sigma_0)^2}} \right) \ , \tag{5.34}$$

where $\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp(-y^2/2) dy$ is the cumulative distribution function of the normal distribution with mean zero and variance one. This result is valid as long as $\sigma_0 > \sigma_{\text{noise}}$. When $\sigma_0 < \sigma_{\text{noise}}$, so that the point reflector is "hidden in noise", then we have $1 - \beta = 1 - \Phi_{\text{TW2}}\big(\Phi_{\text{TW2}}^{-1}(1-\alpha)\big) = \alpha$. Note that, as functions of the number of sensors N, the singular value σ_0 scales as N by Eq. (5.32), while the noise level σ_{noise} scales as 1 (with the Hadamard technique, otherwise it would scale as $\sqrt{N}$). This shows that the detection power increases with the number of sensors.

5.4 Localization and Reconstruction

In this subsection we consider the situation in which the transmitter and receiver arrays are coincident and there are an unknown number r of point

reflectors embedded in the medium. We would like to build algorithms that estimate the number r of point reflectors, estimate their locations $\boldsymbol{z}_j$, and estimate their coefficients of reflection ρ_j (defined by (5.8)). In the first version of the algorithm we assume that the noise level σ_{noise} is known. The algorithm is then the following one.

1. Compute the singular values $\sigma_j(\mathbf{A}^{\text{meas}})$ of the measured array response matrix $\mathbf{A}^{\text{meas}}$.
2. Estimate the number of reflectors by

$$\hat{r} = \max\left\{j, \sigma_j(\mathbf{A}^{\text{meas}}) > r_\alpha \sigma_{\text{noise}}\right\},$$

   where the threshold value r_α, given by (5.33), ensures that the false alarm rate (for the detection of a reflector) is α.
3. For each $j = 1, \ldots, \hat{r}$, estimate the positions $\boldsymbol{z}_j$ of the jth reflector by looking after the position $\hat{\boldsymbol{z}}_j$ of the global maximum of the subspace imaging functional $\mathcal{I}_j(\boldsymbol{z})$ defined by

$$\mathcal{I}_j(\boldsymbol{z}) = \left|\boldsymbol{w}(\boldsymbol{z})^\dagger \boldsymbol{w}_j(\mathbf{A}^{\text{meas}})\right|^2 . \tag{5.35}$$

   Here, $\boldsymbol{w}_j(\mathbf{A}^{\text{meas}})$ is the j-th left singular vector of the measured response matrix (*i.e.*, the left singular vector associated to the j-th largest singular value) and $\boldsymbol{w}(\boldsymbol{z})$ is the normalized vector of Green's functions defined by (5.9).
4. For each $j = 1, \ldots, \hat{r}$, estimate the amplitudes ρ_j of the j-th reflector by

$$\hat{\rho}_j = \Big(\sum_{n=1}^{N} |\hat{G}_0(\omega, \boldsymbol{z}_j, \boldsymbol{x}_n)|^2\Big)^{-1} \hat{\sigma}_j , \tag{5.36}$$

   with $\hat{\sigma}_j$ being the estimator of $\sigma_j(\mathbf{A}^0)$ defined by

$$\hat{\sigma}_j = \frac{\sigma_j(\mathbf{A}^{\text{meas}})}{2} + \Big(\frac{\sigma_j(\mathbf{A}^{\text{meas}})^2}{4} - \sigma_{\text{noise}}^2\Big)^{\frac{1}{2}} . \tag{5.37}$$

The form of the estimator $\hat{\sigma}_j$ comes from the inversion of relation (5.14). If we were using $\sigma_j(\mathbf{A}^{\text{meas}})$ as an estimator of $\sigma_j(\mathbf{A}^0)$, then we would over-estimate the reflectivity coefficients of the reflectors.

Note that we do not need to compute all the singular values of the measured response matrix $\mathbf{A}^{\text{meas}}$, only the singular values larger than $2\sigma_{\text{noise}}$ need to be computed.

If the noise level is not known, the first two steps of the algorithm must be replaced by the following ones:

1. Set $j = 1$ and define $\mathbf{A}^1 = \mathbf{A}^{\text{meas}}$.
2. a) Compute the largest singular value $\sigma_1(\mathbf{A}^j)$ (*i.e.*, the spectral norm of $\mathbf{A}^j$) and the associated singular vectors $\boldsymbol{v}_1(\mathbf{A}^j)$ and $\boldsymbol{w}_1(\mathbf{A}^j)$.

b) Compute the Frobenius norm $\|\mathbf{A}^j\|_F$ and estimate the noise level by

$$\hat{\sigma}_{\text{noise},j} = \left[\frac{\|\mathbf{A}^j\|_F^2 - \sigma_1^2(\mathbf{A}^j)}{N - 4j}\right]^{\frac{1}{2}} . \tag{5.38}$$

c) Compute the test

$$T_j = \begin{cases} 1 & \text{if } \sigma_1(\mathbf{A}^j) > (2 + r_\alpha)\hat{\sigma}_{\text{noise},j} \ , \\ 0 & \text{otherwise,} \end{cases} \tag{5.39}$$

where the threshold value r_α is given by (5.33).

d) If $T_j = 1$ then define $\mathbf{A}^{j+1} = \mathbf{A}^j - \sigma_1(\mathbf{A}^j)\boldsymbol{w}_1(\mathbf{A}^j)\boldsymbol{v}_1(\mathbf{A}^j)^\dagger$, increase j by one, and go to (a).
If $T_j = 0$ then set $\hat{r} = j - 1$ and $\hat{\sigma}_{\text{noise}} = \hat{\sigma}_{\text{noise},j-1}$ (if $j = 1$, then $\hat{\sigma}_{\text{noise}} = \hat{\sigma}_{\text{noise},0} = \|\mathbf{A}^{\text{meas}}\|_F / N^{\frac{1}{2}}$) and go to 3.

The sequence of singular values $\sigma_1(\mathbf{A}^j)$, $j = 1, \ldots, \hat{r}$, is the list of the $\hat{r}$ largest singular values $\sigma_j(\mathbf{A}^{\text{meas}})$ of $\mathbf{A}^{\text{meas}}$. Similarly the sequence of left singular vectors $w_1(\mathbf{A}^j)$, $j = 1, \ldots, \hat{r}$, is the list of the left singular vectors $w_j(\mathbf{A}^{\text{meas}})$ associated to the $\hat{r}$ largest singular values of $\mathbf{A}^{\text{meas}}$. In fact, it is not necessary to compute explicitly the Frobenius norm of $\mathbf{A}^j$ at each step in 2(a). We can compute the Frobenius norm of $\mathbf{A}$ and then use the relation

$$\|\mathbf{A}^j\|_F^2 = \|\mathbf{A}\|_F^2 - \sum_{l=1}^{j-1} \sigma_1^2(\mathbf{A}^l) \ ,$$

or, equivalently, the recursive relation

$$\|\mathbf{A}^1\|_F^2 = \|\mathbf{A}^{\text{meas}}\|_F^2, \qquad \|\mathbf{A}^{j+1}\|_F^2 = \|\mathbf{A}^j\|_F^2 - \sigma_1^2(\mathbf{A}^j), \quad j \geq 1 \ .$$

This algorithm provides an estimator $\hat{\sigma}_{\text{noise}}$ of the noise level σ_{noise} and an estimator $\hat{r}$ of the number r of significant singular values, that is, the number of reflectors. The steps 3 and 4 of the previous algorithm are then used for the localization and characterization of the reflectors, using the estimator $\hat{\sigma}_{\text{noise}}$ for σ_{noise}.

An alternative algorithm to estimate the noise level is based on the minimization of the Kolmogorov-Smirnov distance between the empirical distribution of the (smallest) singular values of the perturbed matrix $\mathbf{A}$ and the theoretical deformed quarter-circle law. This algorithm reduces significantly the bias but it is more computationally intensive. When N is very large, formula (5.38) is sufficient for the noise level estimation.

Instead of $\mathcal{I}_j(\boldsymbol{z})$ defined by (5.35), other subspace imaging functionals such as Kirchhoff-type algorithms can be used. The decomposition of the time-reversal operator (DORT) can also be used [14, 15, 19] for detecting and characterizing the reflectors.

6

Appendix: Basic facts from analysis and probability

In this chapter we review a few results and techniques that are used throughout the book. In Section 6.1 we give a few well-known properties of the Fourier transform. In Section 6.2 we state the divergence theorem. In Section 6.3 we discuss the stationary phase method. Finally Section 6.4 we present some results and tools of probability theory, with a focus on random process modeling.

6.1 Fourier identities

Let $f(t)$ be a "nice" real-valued function (i.e. smooth and integrable). Its Fourier transform is defined by:

$$\hat{f}(\omega) = \int_{\mathbb{R}} f(t) e^{i\omega t} \mathrm{d}t$$

The inverse Fourier transform is:

$$f(t) = \frac{1}{2\pi} \int_{\mathbb{R}} \hat{f}(\omega) e^{-i\omega t} \mathrm{d}\omega.$$

The following table gives the important Fourier identities that are used throughout the book. Their proofs are straightforward by the use of change of variables or integration by parts [20].

$f(t)$	$\hat{f}(\omega)$
$\frac{d^n f}{dt^n}$	$(-i\omega)^n \hat{f}(\omega)$
$f * g(t) = \int f(s) g(t-s) \mathrm{d}s$	$\hat{f}(\omega)\hat{g}(\omega)$
$f(-t)$	$\overline{\hat{f}(\omega)}$
$\int f(s) g(t+s) \mathrm{d}s$	$\overline{\hat{f}(\omega)}\hat{g}(\omega)$

The third identity is useful for time reversal: it expresses the fact that the time reversal operation in the time domain (t) is equivalent to the complex conjugation in the frequency domain (ω). The fourth identity shows that the cross correlation of two signals involves the product of the two Fourier transforms in the frequency domain, one of the transform being complex conjugated.

6.2 Divergence theorem

The divergence theorem is a basic identity that is used in this book. We refer to [20] for a detailed statement and applications. Let V be a bounded open subset of $\mathbb{R}^n$ whose boundary ∂V is $\mathcal{C}^1$. Let $\boldsymbol{f} \in \mathcal{C}^1(\overline{V}, \mathbb{R}^n)$ where $\overline{V} = V \cup \partial V$ is the closure of V. Then

$$\int_V \nabla \cdot \boldsymbol{f}(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} = \int_{\partial V} \boldsymbol{n}(\boldsymbol{x}) \cdot \boldsymbol{f}(\boldsymbol{x}) \mathrm{d}\sigma(\boldsymbol{x}) \,,$$

where $\boldsymbol{n}(\boldsymbol{x})$ is the outward unit normal to ∂V at $\boldsymbol{x} \in \partial V$.

6.3 Stationary-phase method

We review the stationary phase method. For more detail we refer the reader to [38].

The one-dimensional case

Let ϕ and f be two smooth functions from $\mathbb{R}$ to $\mathbb{R}$. Assume that f is compactly supported, that $\phi'(s)$ vanishes only at s_0, and that $f(s_0) \neq 0$, $\phi''(s_0) \neq 0$. The integral

$$I(\varepsilon) = \int_{\mathbb{R}} e^{i \frac{\phi(s)}{\varepsilon}} f(s) \, \mathrm{d}s$$

can be approximated as $\varepsilon \to 0$ by

$$\lim_{\varepsilon \to \infty} \frac{1}{\sqrt{\varepsilon}} I(\varepsilon) e^{-i \frac{\phi(s_0)}{\varepsilon}} = \frac{\sqrt{2\pi}}{\sqrt{|\phi''(s_0)|}} e^{i n^* \frac{\pi}{4}} f(s_0) \,,$$

where $n^* = \mathrm{sgn}(\phi''(s_0))$ [38, Chapter II, section 3].

The n-dimensional case

The stationary phase theorem can be generalized to n-dimensional integrals [38, Chapter IX, Theorem 1]. Let n be a positive integer. Let ϕ and f be two smooth functions from $\mathbb{R}^n$ to $\mathbb{R}$. Assume that f is compactly supported, that

$\nabla\phi(\boldsymbol{s})$ vanishes only at $\boldsymbol{s}_0$, and that $f(\boldsymbol{s}_0)$ and the determinant of the Hessian $\mathbf{H}_{\boldsymbol{s}_0}(\phi)$ of ϕ at $\boldsymbol{s}_0$ are nonzero. The integral

$$I(\varepsilon) = \int_{\mathbb{R}^n} e^{i\frac{\phi(\boldsymbol{s})}{\varepsilon}} f(\boldsymbol{s})\,\mathrm{d}\boldsymbol{s}$$

can be approximated as $\varepsilon \to 0$ by

$$\lim_{\varepsilon\to\infty} \frac{1}{\varepsilon^{n/2}} I(\varepsilon) e^{-i\frac{\phi(\boldsymbol{s}_0)}{\varepsilon}} = \frac{(2\pi)^{n/2}}{\sqrt{|\det \mathbf{H}_{\boldsymbol{s}_0}(\phi)|}} e^{i(2n^*-n)\frac{\pi}{4}} f(\boldsymbol{s}_0)\,, \tag{6.1}$$

where n^* is the number of positive eigenvalues of $\mathbf{H}_{\boldsymbol{s}_0}(\phi)$. The point $\boldsymbol{s}_0$ such that $\nabla\phi(\boldsymbol{s}_0) = 0$ is referred to as the stationary point. In fact there exists a constant C that depends only on f and ϕ such that

$$\left| \frac{1}{\varepsilon^{n/2}} I(\varepsilon) - \frac{(2\pi)^{n/2}}{\sqrt{|\det \mathbf{H}_{\boldsymbol{s}_0}(\phi)|}} e^{i(2n^*-n)\frac{\pi}{4}} f(\boldsymbol{s}_0) e^{i\frac{\phi(\boldsymbol{s}_0)}{\varepsilon}} \right| \leq C\sqrt{\varepsilon}\,. \tag{6.2}$$

A degenerate case

The typical configuration that is encountered in this book is actually degenerate (i.e. there are an infinite number of stationary points and/or the Hessian matrices at the stationary points are not invertible). Degenerate cases can be addressed as explained for instance in [38], Chapter IX, section 4. We give here the result for a particular situation in which there is a line of stationary points and that is of interest for us.

Proposition 6.1. *For any $\varepsilon > 0$, let us consider the integral*

$$I(\varepsilon) = \int_{\mathbb{R}} \int_{\mathbb{R}^n} e^{i\frac{\omega\phi(\boldsymbol{s})}{\varepsilon}} f(\omega) g(\boldsymbol{s})\,\mathrm{d}\boldsymbol{s}\,\mathrm{d}\omega\,,$$

where ϕ and g are smooth functions from $\mathbb{R}^n$ to $\mathbb{R}$, g is compactly supported, f is a smooth function from $\mathbb{R}$ to $\mathbb{R}$ that is compactly supported (away from zero). We assume that $\nabla\phi(\boldsymbol{s})$ vanishes only at $\boldsymbol{s}_0 \in \mathbb{R}^n$ and that the determinant of the Hessian $\mathbf{H}_{\boldsymbol{s}_0}(\phi)$ of ϕ at $\boldsymbol{s}_0$ is nonzero. There are two cases:

1. *If $\phi(\boldsymbol{s}_0) \neq 0$, then*
$$\lim_{\varepsilon\to 0} \frac{1}{\varepsilon^{n/2}} I(\varepsilon) = 0\,. \tag{6.3}$$
2. *If $\phi(\boldsymbol{s}_0) = 0$, then*
$$\lim_{\varepsilon\to 0} \frac{1}{\varepsilon^{n/2}} I(\varepsilon) = \frac{(2\pi)^{n/2} g(\boldsymbol{s}_0)}{\sqrt{|\det \mathbf{H}_{\boldsymbol{s}_0}(\phi)|}} \int e^{i(2n^*-n)\frac{\pi}{4}\mathrm{sgn}(\omega)} \frac{f(\omega)}{|\omega|^{n/2}}\,\mathrm{d}\omega\,, \tag{6.4}$$
where n^* is the number of positive eigenvalues of $\mathbf{H}_{\boldsymbol{s}_0}(\phi)$.

The proof of this proposition is based on the estimate (6.2) that yields for any ω in the support of f (that does not contain zero):

$$\left| \frac{1}{\varepsilon^{n/2}} \int_{\mathbb{R}^n} e^{i\frac{\omega\phi(\boldsymbol{s})}{\varepsilon}} g(\boldsymbol{s})\,\mathrm{d}\boldsymbol{s} - \frac{(2\pi)^{n/2} g(\boldsymbol{s}_0)}{\sqrt{|\det \mathbf{H}_{\boldsymbol{s}_0}(\phi)|}|\omega|^{n/2}} e^{i(2n^*-n)\frac{\pi}{4}\operatorname{sgn}(\omega)} e^{i\frac{\omega\phi(\boldsymbol{s}_0)}{\varepsilon}} \right| \leq C \frac{\sqrt{\varepsilon}}{\sqrt{\omega}},$$

where C depends only on ϕ and g. We can now integrate with respect to ω over the support of f. If $\phi(\boldsymbol{s}_0) \neq 0$, then the integral

$$\int_{\mathbb{R}} e^{i(2n^*-n)\frac{\pi}{4}\operatorname{sgn}(\omega)} e^{i\frac{\omega\phi(\boldsymbol{s}_0)}{\varepsilon}} \frac{f(\omega)}{|\omega|^{n/2}}\,\mathrm{d}\omega$$

goes to zero as $\varepsilon \to 0$ by the Riemann-Lebesgue lemma, which yields (6.3). If $\phi(\boldsymbol{s}_0) = 0$, then we immediately get (6.4).

6.4 Random processes

Here we give some background on random (or stochastic) processes. A reference for this material is [12].

6.4.1 Random variables

A characteristic of noise is that it does not have fixed values in repeated measurements or observations. Let us first consider such a scalar (real-valued) quantity. It can be modeled by a random variable, for which the exact value of a realization is not known, but for which the likelihood or empirical frequency of any measurable set of values can be characterized. The statistical distribution of a random variable can be defined as the probability measure over $\mathbb{R}$ that quantifies the likelihood that the random variable takes values in a particular measurable set. In this section we only address so-called continuous random variables, ie those whose distributions admit densities with respect to the Lebesgue measure over $\mathbb{R}$, as we will never encounter discrete or other singular random variables in the book. The statistical distribution of a random variable can then be characterized by its probability density function (PDF). The PDF of a (real-valued) random variable Z is denoted by $p_Z(z)$:

$$\mathbb{P}\big(Z \in [a,b]\big) = \int_a^b p_Z(z)\,\mathrm{d}z\,.$$

Note that p_Z is a nonnegative function whose total integral is equal to one. Given the PDF it is possible to compute the expectation of a nice function (bounded or positive) of the random variable $\phi(Z)$, which is the weighted average of ϕ with respect to the PDF p_Z:

$$\mathbb{E}[\phi(Z)] = \int_{\mathbb{R}} \phi(z) p_Z(z)\,\mathrm{d}z\,.$$

The most important weighted averages are the first- and second-order moments (we only consider random variables with finite first and second moments in this book). The mean (or expectation) of the random variable Z is defined as

$$\mathbb{E}[Z] = \int_{\mathbb{R}} z p_Z(z) \, \mathrm{d}z \, . \tag{6.5}$$

It is the first-order statistical moment. It is the deterministic value that best approximates the random variable Z in the mean square sense:

$$\mathbb{E}[Z] = \operatorname*{argmin}_{a \in \mathbb{R}} \mathbb{E}\big[(Z-a)^2\big].$$

The variance is defined as

$$\mathrm{Var}(Z) = \mathbb{E}\big[|Z - \mathbb{E}[Z]|^2\big] = \mathbb{E}[Z^2] - \mathbb{E}[Z]^2, \tag{6.6}$$

which is a second-order statistical moment. $\sigma_Z = \sqrt{\mathrm{Var}(Z)}$ is called the standard deviation, which is a measure of the average deviation from the mean.

The PDF of a noise signal is not always known in practical situations. We often use parameters such as mean and variance to describe it. It is then usual to assume that the noise has Gaussian PDF. This can be justified by the maximum of entropy principle, which claims that the PDF that maximizes the entropy $-\int p_Z(z) \ln p_Z(z) \, \mathrm{d}z$ with the constraints $\int p_Z(z) \, \mathrm{d}z = 1$, $\int z p_Z(z) \, \mathrm{d}z = \mu$, and $\int (z-\mu)^2 p_Z(z) \, \mathrm{d}z = \sigma^2$, is the Gaussian PDF

$$p_Z(z) = \frac{1}{\sqrt{2\pi}\sigma} \exp\Big(- \frac{(z-\mu)^2}{2\sigma^2} \Big), \tag{6.7}$$

with mean μ and variance σ^2. If a random variable Z has PDF (6.7), then we write $Z \sim \mathcal{N}(\mu, \sigma^2)$. Moreover, a noise signal often results from the cumulative effect of many uncorrelated sources. As a consequence, based on the central limit theorem, most measurement noise can be treated as Gaussian noise. Recall here the central limit theorem: When a random variable Z is the sum of n independent and identically distributed random variables, then the distribution of Z is a Gaussian distribution with the appropriate mean and variance in the limit $n \to +\infty$, provided the variances are finite.

6.4.2 Random vectors

A d-dimensional random vector $\boldsymbol{Z}$ is a collection of d (real-valued) random variables $(Z_1, \ldots, Z_d)^t$. The distribution of a random vector is characterized by the PDF $p_{\boldsymbol{Z}}$:

$$\mathbb{P}\big(\boldsymbol{Z} \in [a_1, b_1] \times \cdots \times [a_d, b_d]\big) = \int_{[a_1,b_1]\times\cdots[a_d,b_d]} p_{\boldsymbol{Z}}(\boldsymbol{z}) \, \mathrm{d}\boldsymbol{z}, \qquad \text{for all } a_j \le b_j \, .$$

The PDF $p_{\boldsymbol{Z}}$ is a function from $\mathbb{R}^d$ to $[0, \infty)$ whose total integral is equal to one. The random vector $\boldsymbol{Z} = (Z_1, \ldots, Z_d)^t$ is independent if its PDF can be

written as a product of the one-dimensional PDFs of the coordinates of the vector:

$$p_{\boldsymbol{Z}}(\boldsymbol{z}) = \prod_{j=1}^{d} p_{Z_j}(z_j) \text{ for all } \boldsymbol{z} = (z_1, \dots, z_d)^t \in \mathbb{R}^d \,,$$

or equivalently,

$$\mathbb{E}\big[\phi_1(Z_1) \cdots \phi_d(Z_d)\big] = \mathbb{E}\big[\phi_1(Z_1)\big] \cdots \mathbb{E}\big[\phi_d(Z_d)\big],$$

for all continuous bounded functions $\phi_1, \dots, \phi_d \in \mathcal{C}_b(\mathbb{R}, \mathbb{R})$.

Example: a d-dimensional normalized Gaussian random vector $\boldsymbol{Z}$ has the Gaussian PDF

$$p_{\boldsymbol{Z}}(\boldsymbol{z}) = \frac{1}{(2\pi)^{d/2}} \exp\Big(-\frac{|\boldsymbol{z}|^2}{2}\Big) \,.$$

This PDF can be factorized into the product of one-dimensional Gaussian PDFs, which shows that $\boldsymbol{Z}$ is a vector of independent random normalized Gaussian variables $(Z_1, \dots, Z_d)^t$ (normalized means with mean zero and variance one).

As in the case of random variables, we may not always require or may not be able to give a complete statistical description of a random vector. In such cases, we work only with the first and second statistical moments. Let $\boldsymbol{Z} = (Z_i)_{i=1,\dots,d}$ be a random vector. The mean of $\boldsymbol{Z}$ is the vector $\boldsymbol{\mu} = (\mu_j)_{j=1,\dots,d}$:

$$\mu_j = \mathbb{E}[Z_j] \,.$$

The covariance matrix of $\boldsymbol{Z}$ is the matrix $\mathbf{C} = (C_{jl})_{j,l=1,\dots,d}$:

$$C_{jl} = \mathbb{E}\big[(Z_j - \mathbb{E}[Z_j])(Z_l - \mathbb{E}[Z_l])\big] \,.$$

These statistical moments are enough to characterize the first two moments of any linear combination of the components of $\boldsymbol{Z}$. Indeed, if $\boldsymbol{\beta} = (\beta_j)_{j=1,\dots,d} \in \mathbb{R}^d$, then the random variable $Z_{\boldsymbol{\beta}} = \boldsymbol{\beta} \cdot \boldsymbol{Z} = \sum_{j=1}^{d} \beta_j Z_j$ has mean:

$$\mathbb{E}[Z_{\boldsymbol{\beta}}] = \boldsymbol{\beta} \cdot \boldsymbol{\mu} = \sum_{j=1}^{d} \beta_j \mathbb{E}[Z_j] \,,$$

and variance:

$$\mathrm{Var}(Z_{\boldsymbol{\beta}}) = \boldsymbol{\beta}^t \mathbf{C} \boldsymbol{\beta} = \sum_{j,l=1}^{d} C_{jl} \beta_j \beta_l \,.$$

As a byproduct of this result, we can see that the covariance matrix $\mathbf{C}$ is necessarily nonnegative.

If the variables are independent then the covariance matrix is diagonal. In particular:

$$\mathrm{Var}\Big(\sum_{j=1}^{d} Z_j\Big) = \sum_{j=1}^{d} \mathrm{Var}(Z_j) \,.$$

The reciprocal is false in general (*i.e.*, the fact that the covariance matrix is diagonal does not ensure that the vector is independent).

6.4.3 Gaussian random vectors

A Gaussian random vector $\boldsymbol{Z} = (Z_1, \dots, Z_d)$ with mean $\boldsymbol{\mu}$ and covariance matrix $\mathbf{R}$ (write $\boldsymbol{Z} \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{R})$) has the PDF

$$p(\boldsymbol{z}) = \frac{1}{(2\pi)^{d/2}(\det \mathbf{R})^{1/2}} \exp\Big(- \frac{(\boldsymbol{z}-\boldsymbol{\mu})^t \mathbf{R}^{-1}(\boldsymbol{z}-\boldsymbol{\mu})}{2}\Big), \tag{6.8}$$

provided $\mathbf{R}$ is positive. As mentioned in the case of random variables, the Gaussian statistics is the one that is obtained from the maximum of entropy principle (given that the first two moments of the random vector are specified) and from the central limit theorem. This distribution is characterized by the characteristic function or Fourier transform of the PDF:

$$\mathbb{E}[e^{i\boldsymbol{\lambda}\cdot\boldsymbol{Z}}] = \int_{\mathbb{R}^d} e^{i\boldsymbol{\lambda}\cdot\boldsymbol{z}} p(\boldsymbol{z})\,\mathrm{d}\boldsymbol{z} = \exp\Big(i\boldsymbol{\lambda}\cdot\boldsymbol{\mu} - \frac{\boldsymbol{\lambda}^t\mathbf{R}\boldsymbol{\lambda}}{2}\Big), \qquad \boldsymbol{\lambda} \in \mathbb{R}^d, \tag{6.9}$$

which also shows that, if $\boldsymbol{\lambda} \in \mathbb{R}^d$, then the linear combination $\boldsymbol{\lambda}\cdot\boldsymbol{Z}$ is a real-valued Gaussian random variable with mean $\boldsymbol{\lambda}\cdot\boldsymbol{\mu}$ and variance $\boldsymbol{\lambda}^t\mathbf{R}\boldsymbol{\lambda}$.

The expectations of high-order moments of a zero-mean Gaussian vector can be expressed as a sum of second-order moments. For instance, if $\boldsymbol{Z} = (Z_1, Z_2, Z_3, Z_4)$ is a zero-mean Gaussian vector, then

$$\mathbb{E}\Big[\prod_{j=1}^{4} Z_j\Big] = \mathbb{E}\big[Z_1Z_2\big]\mathbb{E}\big[Z_3Z_4\big] + \mathbb{E}\big[Z_1Z_3\big]\mathbb{E}\big[Z_2Z_4\big] + \mathbb{E}\big[Z_1Z_4\big]\mathbb{E}\big[Z_2Z_3\big].$$

6.4.4 Random processes

The perturbations in the index of refraction of an inhomogeneous medium, the wave fluctuations recorded by a receiver array, or the noise that appears in an image are described by functions of space (and/or time) with random values, which are known as random (or stochastic) processes.

Remember that a random variable is a random number, in the sense that a realization of the random variable is a real number and that the statistical distribution of the random variable is characterized by its PDF. In the same way, a random process $(Z(\boldsymbol{x}))_{\boldsymbol{x}\in\mathbb{R}^d}$ is a random function, in the sense that a realization of the random process is a function from $\mathbb{R}^d$ to $\mathbb{R}$, and that the distribution of $(Z(\boldsymbol{x}))_{\boldsymbol{x}\in\mathbb{R}^d}$ is characterized by the finite-dimensional distributions $(Z(\boldsymbol{x}_1), \dots, Z(\boldsymbol{x}_n))^t$, for any n, $\boldsymbol{x}_1, \dots, \boldsymbol{x}_n \in \mathbb{R}^d$ (the fact that the finite-dimensional distributions completely characterize the distribution of the random process is not completely trivial and it follows from Kolmogorov's extension theorem).

As in the case of random variables, we may not always require a complete statistical description of a random process, or we may not be able to obtain it even if desired. In such cases, we work with the first and second statistical moments. The most important ones are

(i) Mean: $\mathbb{E}[Z(\boldsymbol{x})]$;
(ii) Variance: $\mathrm{Var}(Z(\boldsymbol{x})) = \mathbb{E}\big[(Z(\boldsymbol{x}) - \mathbb{E}[Z(\boldsymbol{x})])^2\big]$;
(iii) Covariance function: $R(\boldsymbol{x}, \boldsymbol{x}') = \mathbb{E}\big[(Z(\boldsymbol{x}) - \mathbb{E}[Z(\boldsymbol{x})])(Z(\boldsymbol{x}') - \mathbb{E}[Z(\boldsymbol{x}')])\big]$.

We say that $(Z(\boldsymbol{x}))_{\boldsymbol{x}\in\mathbb{R}^d}$ is a stationary random process if the statistics of the process is invariant to a shift in the origin: for any $\boldsymbol{x}_0 \in \mathbb{R}^d$,

$$(Z(\boldsymbol{x}_0 + \boldsymbol{x}))_{\boldsymbol{x}\in\mathbb{R}^d} \overset{\text{distribution}}{=} (Z(\boldsymbol{x}))_{\boldsymbol{x}\in\mathbb{R}^d}\,.$$

It is a statistical steady state. A necessary and sufficient condition is that, for any integer n, for any $\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_n \in \mathbb{R}^d$, for any bounded continuous function $\phi \in \mathcal{C}_b(\mathbb{R}^n, \mathbb{R})$, we have

$$\mathbb{E}\left[\phi(Z(\boldsymbol{x}_0 + \boldsymbol{x}_1), \ldots, Z(\boldsymbol{x}_0 + \boldsymbol{x}_n))\right] = \mathbb{E}\left[\phi(Z(\boldsymbol{x}_1), \ldots, Z(\boldsymbol{x}_n))\right]\,.$$

6.4.5 Ergodic processes

Let us consider a stationary process such that $\mathbb{E}[|Z(\boldsymbol{x})|] < \infty$. We set $\mu = \mathbb{E}[Z(\boldsymbol{x})]$. The ergodic theorem claims that the time average can be replaced by the statistical average under the so-called ergodic hypothesis [12].

Theorem 6.2. *If $Z(\boldsymbol{x})$ satisfies the ergodic hypothesis, then*

$$\frac{1}{N^d}\int_{[0,N]^d} Z(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \overset{N\to\infty}{\longrightarrow} \mu \qquad \mathbb{P}\text{-almost surely.}$$

The ergodic hypothesis requires that the orbit $(Z(\boldsymbol{x}))_{\boldsymbol{x}\in\mathbb{R}^d}$ visits all of phase space. It is not easy to state and to understand (see Remark 6.4 below), although it seems an intuitive notion. The following example presents an example of a non-ergodic process.

Example 6.3. Let $(Z_1(t))_{t\in\mathbb{R}}$ and $(Z_2(t))_{t\in\mathbb{R}}$ be two ergodic processes (satisfying Theorem 6.2), and denote $\mu_j = \mathbb{E}[Z_j(t)]$, $j = 1, 2$. Assume $\mu_1 \neq \mu_2$. Consider a Bernoulli random variable χ independent of $(Z_1(t))_{t\in\mathbb{R}}$ and $(Z_2(t))_{t\in\mathbb{R}}$ and whose distribution is $\mathbb{P}(\chi = 1) = \mathbb{P}(\chi = 0) = 1/2$. Let $Z(t) = \chi Z_1(t) + (1-\chi)Z_2(t)$, which is a stationary process with mean $\mu = \frac{1}{2}(\mu_1 + \mu_2)$. The time-averaged process satisfies

$$\begin{aligned}\frac{1}{T}\int_0^T Z(t)\mathrm{d}t &= \chi\left(\frac{1}{T}\int_0^T Z_1(t)\mathrm{d}t\right) + (1-\chi)\left(\frac{1}{T}\int_0^T Z_2(t)\mathrm{d}t\right)\\ &\overset{T\to\infty}{\longrightarrow} \chi\mu_1 + (1-\chi)\mu_2,\end{aligned}$$

which is a random limit different from μ. The time-averaged limit depends on χ because Z has been trapped in a part of phase space. The process $(Z(t))_{t\in\mathbb{R}}$ is not ergodic.

Remark 6.4 (Complement on ergodic theory). Here we give a rigorous statement of an ergodic theorem (it is not necessary for what follows). Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space; that is:

- Ω is a non-empty set,
- $\mathcal{A}$ is a σ-algebra on Ω,
- $\mathbb{P} : \mathcal{A} \to [0,1]$ is a probability (i.e. $\mathbb{P}(\Omega) = 1$ and $\mathbb{P}(\cup_j A_j) = \sum_j \mathbb{P}(A_j)$ for any numerable family of disjoint sets $A_j \in \mathcal{A}$).

Let $\theta_{\boldsymbol{x}} : \Omega \to \Omega$, $\boldsymbol{x} \in \mathbb{R}^d$, be a measurable group of shift operators (i.e. $\theta_{\boldsymbol{x}}^{-1}(A) \in \mathcal{A}$ for any $A \in \mathcal{A}$ and $\boldsymbol{x} \in \mathbb{R}^d$, $\theta_{\boldsymbol{0}} = I_d$ and $\theta_{\boldsymbol{x}+\boldsymbol{y}} = \theta_{\boldsymbol{x}} \circ \theta_{\boldsymbol{y}}$ for any $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^d$) that preserves the probability $\mathbb{P}$ (i.e. $\mathbb{P}(\theta_{\boldsymbol{x}}^{-1}(A)) = \mathbb{P}(A)$ for any $A \in \mathcal{A}$ and $\boldsymbol{x} \in \mathbb{R}^d$).

The group $(\theta_{\boldsymbol{x}})_{\boldsymbol{x}\in\mathbb{R}^d}$ is said to be *ergodic* if the invariant sets are negligible or of negligible complementary, i.e.

$$\theta_{\boldsymbol{x}}^{-1}(A) = A \quad \text{for all } \boldsymbol{x} \in \mathbb{R}^d \implies \mathbb{P}(A) = 0 \text{ or } 1.$$

We then have the following proposition.

Proposition 6.5. *Let $f : (\Omega, \mathcal{A}, \mathbb{P}) \to \mathbb{R}$ and $Z(\boldsymbol{x}, \omega) = f(\theta_{\boldsymbol{x}}(\omega))$.*

1. *Z is a stationary random process.*
2. *If $f \in L^1(\mathbb{P})$ and $(\theta_{\boldsymbol{x}})_{\boldsymbol{x}\in\mathbb{R}^d}$ is ergodic, then*

$$\frac{1}{N^d} \int_{[0,N]^d} Z(\boldsymbol{x}, \omega) \mathrm{d}\boldsymbol{x} \overset{N\to\infty}{\longrightarrow} \mathbb{E}[f] = \int_\Omega f \mathrm{d}\mathbb{P} \quad \mathbb{P}\text{-almost surely.}$$

6.4.6 Mean square theory

In this subsection we introduce a weaker form of the ergodic theorem, that holds true under a simple and explicit condition. Let $(Z(\boldsymbol{x}))_{\boldsymbol{x}\in\mathbb{R}^d}$ be a stationary process with finite variance $\mathbb{E}[Z^2(\boldsymbol{0})] < \infty$. We introduce the autocorrelation function

$$c(\boldsymbol{x}) = \mathbb{E}\left[(Z(\boldsymbol{y}) - \mu)(Z(\boldsymbol{y}+\boldsymbol{x}) - \mu)\right],$$

where $\mu = \mathbb{E}[Z(\boldsymbol{y})]$. Both c and μ do not depend on $\boldsymbol{y}$ by stationarity.

By stationarity, c is an even function:

$$\begin{aligned} c(-\boldsymbol{x}) &= \mathbb{E}\left[(Z(\boldsymbol{y}) - \mu)(Z(\boldsymbol{y}-\boldsymbol{x}) - \mu)\right] = \mathbb{E}\left[(Z(\boldsymbol{y}'+\boldsymbol{x}) - \mu)(Z(\boldsymbol{y}') - \mu)\right] \\ &= c(\boldsymbol{x}). \end{aligned}$$

By Cauchy-Schwarz inequality, c reaches its maximum at 0:

$$c(\boldsymbol{x}) \le \mathbb{E}\left[(Z(\boldsymbol{y}) - \mu)^2\right]^{1/2} \mathbb{E}\left[(Z(\boldsymbol{y}+\boldsymbol{x}) - \mu)^2\right]^{1/2} = c(\boldsymbol{0}),$$

and $c(\boldsymbol{0}) = \mathrm{Var}(Z(\boldsymbol{0}))$.

Proposition 6.6. *Assume that* $\int_{\mathbb{R}^d} |c(\boldsymbol{x})| \mathrm{d}\boldsymbol{x} < \infty$. *Let*

$$S(N) = \frac{1}{N^d} \int_{[0,N]^d} Z(\boldsymbol{x}) \mathrm{d}\boldsymbol{x}.$$

Then

$$\mathbb{E}\left[(S(N) - \mu)^2\right] \overset{N\to\infty}{\longrightarrow} 0,$$

more exactly

$$N^d \mathbb{E}\left[(S(N) - \mu)^2\right] \overset{N\to\infty}{\longrightarrow} \int_{\mathbb{R}^d} c(\boldsymbol{x}) \mathrm{d}\boldsymbol{x}.$$

One should interpret the condition $\int_{\mathbb{R}^d} |c(\boldsymbol{x})| \mathrm{d}\boldsymbol{x} < \infty$ as "the autocovariance function $c(\boldsymbol{x})$ decays to 0 sufficiently fast as $|\boldsymbol{x}| \to \infty$." This hypothesis is a mean square version of mixing: $Z(\boldsymbol{y})$ and $Z(\boldsymbol{y} + \boldsymbol{x})$ are approximatively independent for large lags $\boldsymbol{x}$. Mixing substitutes for independence in the law of large numbers. An example of mixing process for $d = 1$ is the piecewise constant process defined by:

$$Z(s) = \sum_{k \in \mathbb{Z}} f_k \mathbf{1}_{[L_k, L_{k+1})}(s),$$

with independent and identically distributed random variables f_k, $L_0 = 0$, $L_k = \sum_{j=1}^{k} l_j$ and independent exponential random variables l_j with mean 1. Here we have $c(\tau) = \mathrm{Var}(f_1) \exp(-|\tau|)$.

Proof. The proof consists in a straightforward calculation. We give it in the case $d = 1$:

$$\begin{aligned}
\mathbb{E}\left[(S(N) - \mu)^2\right] &= \mathbb{E}\left[\frac{1}{N^2} \int_0^N \mathrm{d}t_1 \int_0^N \mathrm{d}t_2 (Z(t_1) - \mu)(Z(t_2) - \mu)\right] \\
&\overset{\text{symmetry}}{=} \frac{2}{N^2} \int_0^N \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}t_2 c(t_1 - t_2) \\
&\overset{\tau = t_1 - t_2}{\underset{h = t_2}{=}} \frac{2}{N^2} \int_0^N \mathrm{d}\tau \int_0^{N-\tau} \mathrm{d}h c(\tau) \\
&= \frac{2}{N^2} \int_0^N \mathrm{d}\tau (N - \tau) c(\tau) = \frac{2}{N} \int_0^\infty \mathrm{d}\tau c_N(\tau),
\end{aligned}$$

where $c_N(\tau) = c(\tau)(1 - \tau/N)\mathbf{1}_{[0,N]}(\tau)$. By Lebesgue's convergence theorem:

$$N\mathbb{E}\left[(S(N) - \mu)^2\right] \overset{N\to\infty}{\longrightarrow} 2 \int_0^\infty c(\tau) \mathrm{d}\tau,$$

which gives the desired result. □

Note that the $L^2(\mathbb{P})$ convergence implies convergence in probability as the limit is deterministic. Indeed, by Chebychev inequality, for any $\delta > 0$,

$$\mathbb{P}\left(|S(N)-\mu|\geq\delta\right)\leq\frac{\mathbb{E}\left[(S(N)-\mu)^2\right]}{\delta^2}\overset{N\to\infty}{\longrightarrow}0.$$

Note also that we can obtain by the same method that, for any $\boldsymbol{k}\in\mathbb{R}^d$,

$$N^d\mathbb{E}\left[\left|\int_{[0,N]^d}\big(Z(\boldsymbol{x})-\mu\big)e^{i\boldsymbol{k}\cdot\boldsymbol{x}}\mathrm{d}\boldsymbol{x}\right|^2\right]\overset{N\to\infty}{\longrightarrow}\int_{\mathbb{R}^d}c(\boldsymbol{x})e^{i\boldsymbol{k}\cdot\boldsymbol{x}}\mathrm{d}\boldsymbol{x},$$

which shows that the Fourier transform of the covariance function of a stationary process is nonnegative. This is a preliminary form of Bochner's theorem which claims that a function $c(\boldsymbol{x})$ is a covariance function of a stationary process if and only if its Fourier transform is nonnegative. The Fourier transform of the covariance function is the power spectral density of the stationary process.

6.4.7 Gaussian processes

We say that a random process $(Z(\boldsymbol{x}))_{\boldsymbol{x}\in\mathbb{R}^d}$ is Gaussian if any linear combination $Z_{\boldsymbol{\lambda}}=\sum_{i=1}^n\lambda_iZ(\boldsymbol{x}_i)$ has Gaussian distribution (for any integer n, $\boldsymbol{x}_i\in\mathbb{R}^d$, $\lambda_i\in\mathbb{R}$). In this case $Z_{\boldsymbol{\lambda}}$ has Gaussian distribution with PDF

$$p_{Z_{\boldsymbol{\lambda}}}(z)=\frac{1}{\sqrt{2\pi}\sigma_{\boldsymbol{\lambda}}}\exp\Big(-\frac{(z-\mu_{\boldsymbol{\lambda}})^2}{2\sigma_{\boldsymbol{\lambda}}^2}\Big),\qquad z\in\mathbb{R}\,,$$

where the mean and variance are given by

$$\mu_{\boldsymbol{\lambda}}=\sum_{i=1}^n\lambda_i\mathbb{E}[Z(\boldsymbol{x}_i)]\,,\qquad\sigma_{\boldsymbol{\lambda}}^2=\sum_{i,j=1}^n\lambda_i\lambda_j\mathbb{E}[Z(\boldsymbol{x}_i)Z(\boldsymbol{x}_j)]-\mu_{\boldsymbol{\lambda}}^2\,.$$

The first two moments of the Gaussian process $(Z(\boldsymbol{x}))_{\boldsymbol{x}\in\mathbb{R}^d}$

$$\begin{aligned}\mu(\boldsymbol{x}_1)&=\mathbb{E}[Z(\boldsymbol{x}_1)],\\ R(\boldsymbol{x}_1,\boldsymbol{x}_2)&=\mathbb{E}\big[(Z(\boldsymbol{x}_1)-\mathbb{E}[Z(\boldsymbol{x}_1)])(Z(\boldsymbol{x}_2)-\mathbb{E}[Z(\boldsymbol{x}_2)])\big],\end{aligned}$$

characterize the finite-dimensional distributions of the process. Indeed, the finite-dimensional distribution of $(Z(\boldsymbol{x}_1),\dots,Z(\boldsymbol{x}_n))^t$ has PDF $p(z_1,\dots,z_n)$ that can be characterized by its Fourier transform:

$$\begin{aligned}&\int_{\mathbb{R}^n}e^{i\sum_{j=1}^n\lambda_jz_j}p(z_1,\dots,z_n)\,\mathrm{d}z_1\cdots\mathrm{d}z_n\\&=\mathbb{E}[e^{i\sum_{j=1}^n\lambda_jZ(\boldsymbol{x}_j)}]=\mathbb{E}[e^{iZ_{\boldsymbol{\lambda}}}]=\int_{\mathbb{R}}e^{iz}p_{Z_{\boldsymbol{\lambda}}}(z)\,dz=\exp\Big(i\mu_{\boldsymbol{\lambda}}-\frac{\sigma_{\boldsymbol{\lambda}}^2}{2}\Big)\\&=\exp\Big(i\sum_{j=1}^n\lambda_j\mu(\boldsymbol{x}_j)-\frac{1}{2}\sum_{j,l=1}^n\lambda_j\lambda_lR(\boldsymbol{x}_j,\boldsymbol{x}_l)\Big)\,,\end{aligned}$$

which shows with (6.9) that $(Z(\boldsymbol{x}_1), \dots, Z(\boldsymbol{x}_n))^t$ has a Gaussian PDF with mean $(\mu(\boldsymbol{x}_j))_{j=1,\dots,n}$ and covariance matrix $(R(\boldsymbol{x}_j, \boldsymbol{x}_l))_{j,l=1,\dots,n}$. As a consequence the distribution of a Gaussian process is characterized by the mean function $(\mu(\boldsymbol{x}_1))_{\boldsymbol{x}_1 \in \mathbb{R}^d}$ and the covariance function $(R(\boldsymbol{x}_1, \boldsymbol{x}_2))_{\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{R}^d}$.

It is rather easy to generate a realization of a Gaussian process $(Z(\boldsymbol{x}))_{\boldsymbol{x} \in \mathbb{R}^d}$ whose mean $\mu(\boldsymbol{x})$ and covariance function $R(\boldsymbol{x}, \boldsymbol{x}')$ are given. If $(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n)$ is a grid of points, then the following algorithm is a random generator of $(Z(\boldsymbol{x}_1), \dots, Z(\boldsymbol{x}_n))^t$:

- compute the mean vector $m_i = \mathbb{E}[Z(\boldsymbol{x}_i)]$ and the covariance matrix $C_{ij} = \mathbb{E}[Z(\boldsymbol{x}_i)Z(\boldsymbol{x}_j)] - \mathbb{E}[Z(\boldsymbol{x}_i)]\mathbb{E}[Z(\boldsymbol{x}_j)]$;
- generate a random vector $\boldsymbol{Y} = (Y_1, \dots, Y_n)^t$ of n independent Gaussian random variables with mean 0 and variance 1 (use randn in matlab, or use the Box-Müller algorithm for instance);
- compute $\boldsymbol{Z} = \boldsymbol{m} + \mathbf{C}^{1/2}\boldsymbol{Y}$.

The vector $\boldsymbol{Z}$ has the distribution of $(Z(\boldsymbol{x}_1), \dots, Z(\boldsymbol{x}_n))^t$ because it has Gaussian distribution (since it is the linear transform of the Gaussian vector $\boldsymbol{Y}$) and it has the desired mean vector and covariance matrix.

Note that the computation of the square root of the matrix $\mathbf{C}$ is expensive from the computational point of view, and one usually chooses to use Cholesky's method to compute it. This simulation method is actually called Cholesky's method. We will see in the next section a faster algorithm when the process is stationary.

6.4.8 Stationary Gaussian processes

We here focus our attention to stationary Gaussian processes. Since the distribution of a Gaussian process is characterized by its first two moments, a Gaussian process is stationary if and only if its mean $\mu(\boldsymbol{x})$ is constant and its covariance function $R(\boldsymbol{x}, \boldsymbol{x}')$ depends only on the lag $\boldsymbol{x}' - \boldsymbol{x}$. Let us consider a stationary Gaussian process $(Z(\boldsymbol{x}))_{\boldsymbol{x} \in \mathbb{R}^d}$ with mean zero and covariance function $c(\boldsymbol{x}) = \mathbb{E}[Z(\boldsymbol{x}')Z(\boldsymbol{x}' + \boldsymbol{x})]$. By Bochner's theorem [26], the Fourier transform of c is necessarily nonnegative. The spectral representation of the real-valued stationary Gaussian process $(Z(\boldsymbol{x}))_{\boldsymbol{x} \in \mathbb{R}^d}$ is:

$$Z(\boldsymbol{x}) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{-i\boldsymbol{k}\cdot\boldsymbol{x}} \sqrt{\hat{c}(\boldsymbol{k})}\hat{n}_{\boldsymbol{k}} \, \mathrm{d}\boldsymbol{k} \,,$$

with $(\hat{n}_{\boldsymbol{k}})_{\boldsymbol{k} \in \mathbb{R}^d}$ a complex white noise, *i.e.*, $\hat{n}_{\boldsymbol{k}}$ is complex-valued, Gaussian, $\hat{n}_{-\boldsymbol{k}} = \overline{\hat{n}_{\boldsymbol{k}}}$, $\mathbb{E}[\hat{n}_{\boldsymbol{k}}] = 0$ and $\mathbb{E}\left[\hat{n}_{\boldsymbol{k}}\overline{\hat{n}_{\boldsymbol{k}'}}\right] = (2\pi)^d \delta(\boldsymbol{k} - \boldsymbol{k}')$ (the representation is formal, one should in fact use stochastic integrals $\mathrm{d}\hat{W}_{\boldsymbol{k}} = \hat{n}_{\boldsymbol{k}}\mathrm{d}\boldsymbol{k}$ with respect to Brownian motions). A complex white noise is actually the Fourier transform of a real white noise: we have $\hat{n}_{\boldsymbol{k}} = \int e^{i\boldsymbol{k}\cdot\boldsymbol{x}} n(\boldsymbol{x})\mathrm{d}\boldsymbol{x}$ where $(n(\boldsymbol{x}))_{\boldsymbol{x} \in \mathbb{R}^d}$ is a real white noise, i.e $n(\boldsymbol{x})$ real-valued, Gaussian, $\mathbb{E}[n(\boldsymbol{x})] = 0$, and $\mathbb{E}[n(\boldsymbol{x})n(\boldsymbol{x}')] = \delta(\boldsymbol{x} - \boldsymbol{x}')$.

It is quite easy to generate a realization of a stationary Gaussian process (with mean zero and covariance function $c(\boldsymbol{x})$) using its spectral representation and Fast Fourier Transforms. In dimension $d = 1$, if we fix a grid of points $x_j = (j-1)\Delta x$, $j = 1, \ldots, n$, then one can generate a realization of the random vector $(Z(x_1), \ldots, Z(x_n))^t$ by the following algorithm:

- evaluate the covariance vector $\boldsymbol{c} = (c(x_1), \ldots, c(x_n))^t$;
- generate a random vector $\boldsymbol{Y} = (Y_1, \ldots, Y_n)^t$ of n independent Gaussian random variables with mean 0 and variance 1;
- filter with the square root of the discrete Fourier transform (DFT) of $\boldsymbol{c}$:

$$\boldsymbol{Z} = \mathrm{IFT}\big(\sqrt{\mathrm{DFT}(\boldsymbol{c})} \cdot \times \mathrm{DFT}(\boldsymbol{Y})\big)\,,$$

  where $\cdot\times$ is the element-wise multiplication.

Then the vector $\boldsymbol{Z}$ is a realization of $(Z(x_1), \ldots, Z(x_n))^t$. In practice one uses FFT and IFFT instead of DFT and IFT, and one obtains a periodized version of the random vector $(Z(x_1), \ldots, Z(x_n))^t$, due to the FFT. This is good enough when the size $n\Delta x$ is much larger than the correlation length of the process (*i.e.* the width of the covariance function c). It is possible to remove the end points of the grid over a band of thickness of the order of the correlation length to remove this periodization effect. In practice this spectral algorithm is more efficient than the Cholesky's method. It can be readily extended to generate a realization of a stationary Gaussian process in dimension $d > 1$.

6.4.9 Vector- and complex-valued Gaussian processes

We finally introduce Gaussian multi-valued processes, which are natural extensions of the real-valued Gaussian processes discussed in the previous subsections.

We say that a $\mathbb{R}^p$-valued process $(\boldsymbol{Z}(\boldsymbol{x}))_{\boldsymbol{x}\in\mathbb{R}^d}$ is a Gaussian process if any finite linear combination $\sum_i \lambda_i Z_{j_i}(\boldsymbol{x}_i)$ is a real-valued Gaussian random variable, for $\lambda_i \in \mathbb{R}$, $j_i \in \{1, \ldots, p\}$, $\boldsymbol{x}_i \in \mathbb{R}^d$. Therefore the coordinate functions $(Z_1(\boldsymbol{x}))_{\boldsymbol{x}\in\mathbb{R}^d}$, …, $(Z_p(\boldsymbol{x}))_{\boldsymbol{x}\in\mathbb{R}^d}$ are real-valued random processes, more exactly they are correlated real-valued Gaussian processes. The distribution of the $\mathbb{R}^p$-valued Gaussian process $(\boldsymbol{Z}(\boldsymbol{x}))_{\boldsymbol{x}\in\mathbb{R}^d}$ is characterized by its vector-valued mean function $\boldsymbol{\mu}(\boldsymbol{x}) = \mathbb{E}[\boldsymbol{Z}(\boldsymbol{x})]$ and its matrix-valued covariance function $\mathbf{R}(\boldsymbol{x}, \boldsymbol{x}') = (R_{ij}(\boldsymbol{x}, \boldsymbol{x}'))_{i,j=1,\ldots,p}$, with $R_{ij}(\boldsymbol{x}, \boldsymbol{x}') = \mathbb{E}[Z_i(\boldsymbol{x})Z_j(\boldsymbol{x}')]$. In particular, the coordinate functions $(Z_i(\boldsymbol{x}))_{\boldsymbol{x}\in\mathbb{R}^d}$ and $(Z_j(\boldsymbol{x}))_{\boldsymbol{x}\in\mathbb{R}^d}$ are independent if and only if $R_{ij}(\boldsymbol{x}, \boldsymbol{x}') = 0$ for all $\boldsymbol{x}, \boldsymbol{x}' \in \mathbb{R}^d$.

We say that a $\mathbb{C}$-valued process $(Z(\boldsymbol{x}))_{\boldsymbol{x}\in\mathbb{R}^d}$ is a Gaussian process if any finite linear combination $\sum_i \lambda_i \mathrm{Re}(Z(\boldsymbol{x}_i)) + \sum_j \lambda'_j \mathrm{Im}(Z(\boldsymbol{x}'_j))$ is a real-valued Gaussian random variable. A $\mathbb{C}$-valued Gaussian process $(Z(\boldsymbol{x}))_{\boldsymbol{x}\in\mathbb{R}^d}$ can be seen as a $\mathbb{R}^2$-valued Gaussian process $(\tilde{\boldsymbol{Z}}(\boldsymbol{x}))_{\boldsymbol{x}\in\mathbb{R}^d}$ with $\tilde{\boldsymbol{Z}} = (\mathrm{Re}(Z), \mathrm{Im}(Z))^t$. Its distribution can be characterized by the vector-valued mean function $\tilde{\boldsymbol{\mu}}(\boldsymbol{x})$

and the matrix-valued covariance function $\tilde{\mathbf{R}}$ associated to $(\tilde{\boldsymbol{Z}}(\boldsymbol{x}))_{\boldsymbol{x}\in\mathbb{R}^d}$. It can as well be characterized by the complex-valued mean function $\mu(\boldsymbol{x}) = \mathbb{E}[Z(\boldsymbol{x})]$, the covariance function $R(\boldsymbol{x},\boldsymbol{x}') = \mathbb{E}[(Z(\boldsymbol{x}) - \mu(\boldsymbol{x}))(\overline{Z(\boldsymbol{x}')} - \overline{\mu(\boldsymbol{x}')})]$, and the relation function $Q(\boldsymbol{x},\boldsymbol{x}') = \mathbb{E}[(Z(\boldsymbol{x}) - \mu(\boldsymbol{x}))(Z(\boldsymbol{x}') - \mu(\boldsymbol{x}'))]$. The PDF of the random vector $(Z(\boldsymbol{x}_1),\dots,Z(\boldsymbol{x}_n))$ (with respect to the Lebesgue measure over $\mathbb{C}^n$) is

$$p(\boldsymbol{z}) = \frac{1}{\pi^n \det(\mathbf{D})^{1/2} \det(\overline{\mathbf{D}} - \overline{\mathbf{C}}^t \mathbf{D}^{-1}\mathbf{C})^{1/2}} \times \exp\left[-\frac{1}{2} \begin{pmatrix} \overline{\boldsymbol{z}} - \overline{\boldsymbol{m}} \\ \boldsymbol{z} - \boldsymbol{m} \end{pmatrix}^t \begin{pmatrix} \mathbf{D} & \mathbf{C} \\ \overline{\mathbf{C}}^t & \overline{\mathbf{D}} \end{pmatrix}^{-1} \begin{pmatrix} \boldsymbol{z} - \boldsymbol{m} \\ \overline{\boldsymbol{z}} - \overline{\boldsymbol{m}} \end{pmatrix} \right],$$

where $D_{ij} = R(\boldsymbol{x}_i,\boldsymbol{x}_j)$, $C_{ij} = Q(\boldsymbol{x}_i,\boldsymbol{x}_j)$, $m_i = \mu(\boldsymbol{x}_i)$.

A circularly symmetric complex Gaussian process is a $\mathbb{C}$-valued Gaussian process such that $\mu(\boldsymbol{x}) = 0$ and $Q(\boldsymbol{x},\boldsymbol{x}') = 0$ for any $\boldsymbol{x},\boldsymbol{x}' \in \mathbb{R}^d$. Its distribution is characterized by its covariance function $R(\boldsymbol{x},\boldsymbol{x}') = \mathbb{E}[Z(\boldsymbol{x})\overline{Z(\boldsymbol{x}')}]$. If, additionally, the covariance function R is real-valued, then the real and imaginary parts $(\mathrm{Re}(Z(\boldsymbol{x})))_{\boldsymbol{x}\in\mathbb{R}^d}$ and $(\mathrm{Im}(Z(\boldsymbol{x})))_{\boldsymbol{x}\in\mathbb{R}^d}$ are independent and identically distributed, they are both Gaussian processes with mean zero and covariance function $R(\boldsymbol{x},\boldsymbol{x}')/2$.

6.5 Some results on Gaussian random fields

Let $(Z(\boldsymbol{x}))_{\boldsymbol{x}\in\mathbb{R}^3}$ be a stationary Gaussian random field with mean zero. The statistical distribution of the random field is characterized by the covariance function:

$$c(\boldsymbol{x}) = \mathbb{E}\big[Z(\boldsymbol{x}')Z(\boldsymbol{x}'+\boldsymbol{x})\big].$$

We assume here that c is smooth, so that the realizations of the random field are smooth [3, Theorem 1.4.2]. As we will see below, the relevant statistical information about local and global maxima of the field is in the variance

$$u_0^2 = c(\mathbf{0}) = \mathbb{E}[Z(\boldsymbol{x})^2]$$

and in the Hessian matrix of the correlation function

$$\boldsymbol{\Lambda} = \Big(\frac{\mathbb{E}\big[\partial_{x_j} Z(\boldsymbol{x}) \partial_{x_l} Z(\boldsymbol{x})\big]}{\mathbb{E}\big[Z(\boldsymbol{x})^2\big]} \Big)_{j,l=1,\dots,3} = \Big(-\frac{\partial^2_{x_j x_l} c(\mathbf{0})}{c(\mathbf{0})} \Big)_{j,l=1,\dots,3}.$$

6.5.1 Local maxima of a Gaussian random field

Let $\Omega \subset \mathbb{R}^3$ be a bounded domain. Let us denote by M_u^Ω the number of local maxima of $Z(\boldsymbol{x})$ in Ω with values larger than u:

$$M_u^\Omega = \mathrm{Card}\big\{ \text{ local maxima of } (Z(\boldsymbol{x}))_{\boldsymbol{x}\in\Omega} \text{ with values larger than } u \ \big\}.$$

We have [2, Theorem 6.3.1]

$$\mathbb{E}[M_u^\Omega] = \frac{|\Omega|}{V_c}\frac{u^2}{u_0^2}\exp\big(-\frac{u^2}{2u_0^2}\big)\Big(1+O\Big(\frac{u_0}{u}\Big)\Big), \text{ for } u \gg u_0,$$

where V_c is the hotspot volume defined in terms of the determinant of the Hessian of the correlation function as:

$$V_c = \frac{4\pi^2}{(\det \boldsymbol{\Lambda})^{1/2}}.$$

6.5.2 Global maximum of a Gaussian random field

Let us denote by $Z_{\max}^\Omega$ the global maximum of the field over the domain Ω:

$$Z_{\max}^\Omega = \max_{\boldsymbol{x}\in\Omega} Z(\boldsymbol{x}).$$

Using [2, Theorem 6.9.4] when $|\Omega| \gg V_c$, the statistical distribution of $Z_{\max}^\Omega$ is of the form

$$Z_{\max}^\Omega = u_0\Big[A\Big(\frac{|\Omega|}{V_c}\Big) + B\Big(\frac{|\Omega|}{V_c}\Big)Z_0\Big],$$

where Z_0 follows a Gumbel distribution with cumulative distribution function $\mathbb{P}(Z_0 \le x) = \exp(-e^{-x})$,

$$A(V) = \sqrt{2\log(V)} + \frac{\log\big[2\log(V)\big]}{\sqrt{2\log(V)}},$$
$$B(V) = \frac{1}{\sqrt{2\log(V)}}.$$

To leading order, the value of the global maximum is deterministic and given by

$$Z_{\max}^\Omega \simeq u_0\sqrt{2\log\Big(\frac{|\Omega|}{V_c}\Big)}.$$

6.5.3 The local shape of a local maximum

We first state a classical and fundamental lemma about Gaussian vectors.

Lemma 6.7. *Let us consider a $\mathbb{R}^{n_1+n_2}$-valued random vector $\begin{pmatrix}\boldsymbol{y}_1\\ \boldsymbol{y}_2\end{pmatrix}$ with Gaussian statistics:*

$$\mathcal{L}\left(\begin{pmatrix}\boldsymbol{y}_1\\ \boldsymbol{y}_2\end{pmatrix}\right) \sim \mathcal{N}\left(\begin{pmatrix}\overline{\boldsymbol{y}}_1\\ \overline{\boldsymbol{y}}_2\end{pmatrix}, \begin{pmatrix}\mathbf{R}_{11} & \mathbf{R}_{12}\\ \mathbf{R}_{21} & \mathbf{R}_{22}\end{pmatrix}\right).$$

The mean vectors $\overline{\boldsymbol{y}}_1$ and $\overline{\boldsymbol{y}}_2$ are in $\mathbb{R}^{n_1}$ and $\mathbb{R}^{n_2}$, respectively, the covariance matrix $\mathbf{R}_{11}$ has size $n_1\times n_1$, $\mathbf{R}_{12}$ has size $n_1\times n_2$, $\mathbf{R}_{21} = \mathbf{R}_{12}^T$ has size $n_2\times n_1$,

and $\mathbf{R}_{22}$ *has size* $n_2 \times n_2$. *Assume that the distribution of* $\boldsymbol{y}_2$ *is not degenerate, i.e., that* $\mathbf{R}_{22}$ *is invertible. Then, conditioned on* $\boldsymbol{y}_2$, *the distribution of* $\boldsymbol{y}_1$ *is Gaussian:*

$$\mathcal{L}\Big(\boldsymbol{y}_1|\boldsymbol{y}_2\Big) \sim \mathcal{N}\Big(\overline{\boldsymbol{y}}_1 + \mathbf{R}_{12}\mathbf{R}_{22}^{-1}(\boldsymbol{y}_2 - \overline{\boldsymbol{y}}_2), \mathbf{R}_{11} - \mathbf{R}_{12}\mathbf{R}_{22}^{-1}\mathbf{R}_{21}\Big).$$

Using this lemma one can show that, given that the random field $Z(\boldsymbol{x})$ has a local maximum at $\boldsymbol{x}_0$ with peak amplitude u (with $u \gg u_0$), then we have locally around $\boldsymbol{x}_0$:

$$Z(\boldsymbol{x}) \simeq u\Big[\frac{c(\boldsymbol{x} - \boldsymbol{x}_0)}{u_0^2} + o(1)\Big], \qquad [u \gg u_0].$$

Proof. Let us fix $\boldsymbol{x}_0 \in \mathbb{R}^3$. The random vector

$$\begin{pmatrix} Z(\boldsymbol{x}) \\ Z(\boldsymbol{x}_0) \\ \nabla Z(\boldsymbol{x}_0) \end{pmatrix}$$

has Gaussian distribution:

$$\mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} c(\mathbf{0}) & c(\boldsymbol{x} - \boldsymbol{x}_0) & -\nabla c(\boldsymbol{x} - \boldsymbol{x}_0)^T \\ c(\boldsymbol{x} - \boldsymbol{x}_0) & c(\mathbf{0}) & 0 \\ -\nabla c(\boldsymbol{x} - \boldsymbol{x}_0) & 0 & c(\mathbf{0})\boldsymbol{\Lambda} \end{pmatrix}\right).$$

Applying the above lemma, the distribution of $Z(\boldsymbol{x})$ given $Z(\boldsymbol{x}_0) = u$ and $\nabla Z(\boldsymbol{x}_0) = \mathbf{0}$ is

$$\mathcal{N}\Big(\frac{c(\boldsymbol{x} - \boldsymbol{x}_0)}{c(\mathbf{0})}u, c(\mathbf{0}) - \frac{\nabla c(\boldsymbol{x} - \boldsymbol{x}_0)^T \boldsymbol{\Lambda}^{-1} \nabla c(\boldsymbol{x} - \boldsymbol{x}_0)}{c(\mathbf{0})}\Big).$$

This gives the desired result when $u \gg u_0 = c(\mathbf{0})^{1/2}$.

6.6 Random Matrix Theory

Our main goal in this appendix is to describe the distribution of the eigenvalues or singular values of random matrices. Here we analyze the eigenvalues of classical models of random diagonalizable matrices and we then extend the results to the singular values of the random matrices of interest in the context of imaging in Chapter 5. We look for different types of results. On the one hand we look for the description of the global distribution of the singular values, and on the other hand we also look for a detailed description of the maximal singular value. There are different types of approaches. Some of them are based on the asymptotic expansions of explicit expressions (for some special models) and they give the more detailed results. Other approaches are based on tools of complex analysis (using in particular Stieltjes transforms of

measures) and allow to obtain asymptotic results for a large class of random matrices. As will be shown, the distribution of the singular values depends on the structure of the matrix (symmetric, Hermitian, etc), on the correlation between the random coefficients of the matrix, but not much on the marginal distribution of the coefficients (in the limit of large matrices). A remarkable point is that there are universal results, corresponding to a kind of "central limit theorem", but with unusual scaling and limit distribution. A good reference is the book [34].

6.6.1 Gaussian Orthogonal Ensemble

We consider first a simple model of random matrices in

$$\mathcal{S}_n = \{ \text{ symmetric real matrices of size } n \times n \} .$$

Consider the random matrix $\mathbf{M}$ such that

(i) $(M_{jl})_{1 \le j < l \le n}$ are independent and identically distributed $\mathcal{N}(0,1)$;
(ii) $(M_{jj})_{1 \le j \le n}$ are independent and identically distributed $\mathcal{N}(0,2)$;
(iii) $M_{jl} = M_{lj}$ for $1 \le j < l \le n$.

We then say that $\mathbf{M}$ is a GOE(n)-matrix (GOE stands for Gaussian Orthogonal Ensemble). This model is invariant by any rotation: If $\mathbf{P}$ is a $n \times n$ orthogonal matrix (*i.e.*, $\mathbf{P}^T\mathbf{P} = \mathbf{P}\mathbf{P}^T = \mathbf{I}$) and $\mathbf{M}$ is a GOE(n)-matrix, then the random matrix $\widetilde{\mathbf{M}} = \mathbf{P}\mathbf{M}\mathbf{P}^T$ is a GOE(n)-matrix. Indeed, the matrix $\widetilde{\mathbf{M}}$ is still symmetric, with Gaussian entries, and one can check that the means and covariances of the entries are the ones of the GOE(n)-matrix distribution.

By independence of the coefficients the distribution of $\mathbf{M}$ has a pdf with respect to the measure $d\mathbf{M} = \prod_{1 \le j < l \le n} dM_{jl} \prod_{1 \le j \le n} dM_{jj}$:

$$p_{\mathcal{S}}(\mathbf{M}) = \prod_{1 \le j < l \le n} \frac{1}{\sqrt{2\pi}} \exp\Big(-\frac{M_{jl}^2}{2} \Big) \prod_{1 \le j \le n} \frac{1}{\sqrt{4\pi}} \exp\Big(-\frac{M_{jj}^2}{4} \Big) .$$

Using the symmetry $M_{jl} = M_{lj}$ we can write the pdf as

$$p_{\mathcal{S}}(\mathbf{M}) = c_n \exp\Big(-\frac{\sum_{1 \le i,j \le n} M_{ij}^2}{4} \Big) ,$$

with $c_n = (2\pi)^{-n(n-1)/2} (4\pi)^{-n/2}$. We recognize the trace of $\mathbf{M}\mathbf{M}^t = \mathbf{M}^2$ in the exponent which allows us to write

$$p_{\mathcal{S}}(\mathbf{M}) = c_n \exp\Big(-\frac{\mathrm{Tr}(\mathbf{M}^2)}{4} \Big) . \tag{6.10}$$

The following theorem holds.

Theorem 6.8. *Let* $\mathbf{M}$ *be a GOE*(n)*-matrix. Let* $(\lambda_1, \dots, \lambda_n)$ *be the eigenvalues of* M *sorted in decreasing order. Then the distribution of the random vector* $(\lambda_1, \dots, \lambda_n)$ *has the pdf*

$$p(\lambda_1, \dots, \lambda_n) = \frac{1}{Z_n} \exp\Big(-\frac{1}{4} \sum_{i=1}^{n} \lambda_i^2 \Big) \prod_{1 \le i < j \le n} (\lambda_i - \lambda_j) \, ,$$

for some normalizing constant Z_n*.*

We can notice that the eigenvalues are not independent because the pdf cannot be written as a product of one-dimensional pdfs. In particular, the presence of the product $\prod_{1 \le i < j \le n} (\lambda_i - \lambda_j)$ is a manifestation of level repulsion: the pdf becomes small when two eigenvalues are close to each other. Note finally that this theorem is valid for any n and gives the complete answer to the question: what is the distribution of the eigenvalues of a GOE(n) matrix ? However the answer takes a form that is not easy to understand when n is large, and it is desirable to determine the behavior of the eigenvalues in this asymptotic framework.

Proof (of Theorem 6.8). We give the essential steps of the proof. For any symmetric and real matrix $\mathbf{M}$ we can associate the eigenvalues $(\lambda_j(\mathbf{M}))_{j=1,\dots,n}$. This map is well-defined provided we require the eigenvalues to be sorted in decreasing order. From the spectral representation of symmetric matrices, the matrix $\mathbf{M}$ can be written as

$$\mathbf{M} = \mathbf{P} \boldsymbol{\Lambda} \mathbf{P}^T \, ,$$

where $\mathbf{P}$ is an orthogonal matrix and $\boldsymbol{\Lambda}$ is the diagonal matrix with diagonal entries $\lambda_l(\mathbf{M})$, $l = 1, \dots, n$. For each $l = 1, \dots, n$, the vector $(P_{jl})_{j=1,\dots,n}$ is an eigenvector of $\mathbf{M}$ with eigenvalue λ_l, and we can choose the first nonzero coefficient of the vector $(P_{jl})_{j=1,\dots,n}$ to be positive. Note that the decomposition is then unique provided the eigenvalues are simple, but it is not unique if there are multiple eigenvalues.

If we denote by $d\mathbf{P}$ the Haar measure on the set $\mathcal{O}_n$ of orthogonal matrices (which is the uniform measure on $\mathcal{O}_n$) and by $d\boldsymbol{\Lambda}$ the measure $d\lambda_1 \cdots d\lambda_n$, then for any test function f

$$\begin{aligned}
\mathbb{E}[f(\lambda_1, \dots, \lambda_n)] &= \int f(\lambda_1(\mathbf{M}), \dots, \lambda_n(\mathbf{M})) p_{\mathcal{S}}(\mathbf{M}) d\mathbf{M} \\
&= \iint f(\lambda_1, \dots, \lambda_n) p_{\mathcal{S}}(\mathbf{P} \boldsymbol{\Lambda} \mathbf{P}^T) \Big| \mathrm{Jac} \frac{\partial \mathbf{M}}{\partial \boldsymbol{\Lambda} \partial \mathbf{P}} \Big| d\mathbf{P} d\boldsymbol{\Lambda} \\
&= \iint f(\lambda_1, \dots, \lambda_n) c_n e^{-\frac{1}{4} \sum_{j=1}^n \lambda_j^2} \Big| \mathrm{Jac} \frac{\partial \mathbf{M}}{\partial \boldsymbol{\Lambda} \partial \mathbf{P}} \Big| d\mathbf{P} d\boldsymbol{\Lambda} \, ,
\end{aligned}$$

where we have used the fact that

$$\mathrm{Tr}((\mathbf{P} \boldsymbol{\Lambda} \mathbf{P}^T)^2) = \mathrm{Tr}(\mathbf{P} \boldsymbol{\Lambda} \mathbf{P}^T \mathbf{P} \boldsymbol{\Lambda} \mathbf{P}^T) = \mathrm{Tr}(\boldsymbol{\Lambda}^2) = \sum_{j=1}^{n} \lambda_j^2 \, .$$

The set $\mathcal{O}_n$ of orthogonal matrices is a set with $n(n-1)/2$ parameters (it is a compact Lie group of dimension $n(n-1)/2$). Therefore the Jacobian matrix can be written as

$$\mathrm{Jac}\frac{\partial \mathbf{M}}{\partial \boldsymbol{\Lambda} \partial \mathbf{P}} = \begin{pmatrix} \frac{\partial \mathbf{M}}{\partial \boldsymbol{\Lambda}} & n \text{ lines} \\ \frac{\partial \mathbf{M}}{\partial \mathbf{P}} & n(n-1)/2 \text{ lines} \end{pmatrix} .$$

Note that $\frac{\partial \mathbf{M}}{\partial \boldsymbol{\Lambda}}$ does not depend on $(\lambda_j)_{j=1}^n$ and $\frac{\partial \mathbf{M}}{\partial \mathbf{P}}$ is linear in $(\lambda_j)_{j=1}^n$, because the decomposition $\mathbf{M} = \mathbf{P}\boldsymbol{\Lambda}\mathbf{P}^T$ is linear in $\boldsymbol{\Lambda}$. The determinant of the Jacobian matrix is therefore a polynomial in $(\lambda_j)_{j=1}^n$ of degree $n(n-1)/2$. Moreover, this polynomial cancels as soon as $\lambda_j = \lambda_l$ for some pair of indices (j,l) (because the map $(\boldsymbol{\Lambda}, \mathbf{P}) \mapsto \mathbf{M}$ is not one-to-one), therefore it must contain the factor $\prod_{j<l}(\lambda_j - \lambda_l)$. Since the degree of the polynomial is $n(n-1)/2$, it must be proportional to this product and we must have

$$\left|\mathrm{Jac}\frac{\partial \mathbf{M}}{\partial \boldsymbol{\Lambda} \partial \mathbf{P}}\right| = \prod_{j<l}(\lambda_j - \lambda_l) F(\mathbf{P})$$

for some function F that we do not need to compute. We have

$$\mathbb{E}[f(\lambda_1, \ldots, \lambda_n)] = \int f(\lambda_1, \ldots, \lambda_n) e^{-\frac{1}{4}\sum_{j=1}^n \lambda_j^2} \prod_{j<l}(\lambda_j - \lambda_l) d\boldsymbol{\Lambda} \times \int c_n F(\mathbf{P}) d\mathbf{P} ,$$

which gives the desired result. □

6.6.2 Gaussian Unitary Ensemble

The previous result can be extended to the Gaussian Unitary Ensemble (GUE). This is a model of random matrices in

$$\mathcal{U}_n = \{ \text{ Hermitian complex matrices of size } n \times n \} .$$

Consider the random matrix $\mathbf{M}$ such that

(i) $(M_{jl} = M_{jl}^R + iM_{jl}^I)_{1 \le j < l \le n}$ where M_{jl}^R and M_{jl}^I are independent and identically distributed $\mathcal{N}(0, 1/2)$;
(ii) $(M_{jj})_{1 \le j \le n}$ are independent and identically distributed $\mathcal{N}(0,1)$;
(iii) $M_{jl} = \overline{M_{lj}}$ for $1 \le j < l \le n$.

We then say that $\mathbf{M}$ is a GUE(n)-matrix. This model is invariant by any unitary transform: If $\mathbf{P}$ is a $n \times n$ unitary matrix and $\mathbf{M}$ is a GUE(n)-matrix, then the random matrix $\widetilde{\mathbf{M}} = \mathbf{P}\mathbf{M}\mathbf{P}^\dagger$ is a GUE(n)-matrix, where † stands for the conjugate transpose. Indeed, the matrix $\widetilde{\mathbf{M}}$ is still Hermitian, with Gaussian entries, and one can check that the means and covariances of the entries are the ones of the GUE(n)-matrix distribution. By independence of the coefficients the distribution of $\mathbf{M}$ has a pdf with respect to the measure $d\mathbf{M} = \prod_{1 \le j < l \le n} dM_{jl}^R dM_{jl}^I \prod_{1 \le j \le n} dM_{jj}$ which is of the form

$$p_{\mathcal{U}}(\mathbf{M}) = c_n \exp\Big(-\frac{\mathrm{Tr}(\mathbf{M}^2)}{2}\Big)$$

for some constant c_n. Let $(\lambda_1, \ldots, \lambda_n)$ be the eigenvalues of $\mathbf{M}$ sorted in decreasing order. Then the distribution of the random vector $(\lambda_1, \ldots, \lambda_n)$ has the pdf

$$p(\lambda_1, \ldots, \lambda_n) = \frac{1}{Z_n} \exp\Big(-\frac{1}{2}\sum_{i=1}^{n} \lambda_i^2\Big) \prod_{1 \le i < j \le n} (\lambda_i - \lambda_j)^2$$

for some normalizing constant Z_n.

6.6.3 Asymptotic distributions of the eigenvalues

Given the complete expressions of the joint pdf of the eigenvalues $(\lambda_1, \ldots, \lambda_n)$ of a GOE(n)-matrix or of a GUE(n)-matrix, it is possible to study their asymptotic distributions as $n \to +\infty$ [34]. We first need to take an appropriate scaling. We will analyze the asymptotic behavior of the eigenvalues $(\lambda_1^{(n)}, \ldots, \lambda_n^{(n)})$ of the matrix $\mathbf{M}/\sqrt{n}$, where $\mathbf{M}$ is a GOE(n)-matrix or a GUE(n)-matrix. Let us first consider the density of states, that is, the random counting measure $\rho^{(n)}$ such that

$$\rho^{(n)}([a,b]) = \frac{1}{n}\mathrm{Card}\Big\{j = 1, \ldots, n\,,\ \lambda_j^{(n)} \in [a,b]\Big\} \quad \text{for any } a < b\,.$$

The random quantity $\rho^{(n)}([a,b])$ is the proportion of eigenvalues of the random matrix $\mathbf{M}/\sqrt{n}$ that belong to the interval $[a,b]$. It can be shown that the density of states converges with probability one as $n \to \infty$ to a deterministic continuous measure ρ_{sc}

$$\rho^{(n)}([a,b]) \overset{n\to\infty}{\longrightarrow} \int_a^b \rho_{\mathrm{sc}}(\lambda) d\lambda\,,$$

where

$$\rho_{\mathrm{sc}}(\lambda) = \frac{1}{2\pi}\sqrt{4-\lambda^2}\mathbf{1}_{[-2,2]}(\lambda)\,,$$

and $\mathbf{1}_{[-2,2]}$ is the characteristic function of $[-2,2]$. The distribution with density $\lambda \mapsto (1/2\pi)\sqrt{4-\lambda^2}\mathbf{1}_{[-2,2]}(\lambda)$ is called the semi-circle distribution as its graph has the form of a semi-circle. Here, $\rho_{\mathrm{sc}}(\lambda)d\lambda$ gives the proportion of eigenvalues in the interval $[\lambda, \lambda + d\lambda]$. We find that the asymptotic density of states is the pdf of the semi-circle distribution $\rho_{\mathrm{sc}}(\lambda)$.

Let us now consider the largest eigenvalue

$$\lambda_{\max}^{(n)} = \max\big(\lambda_1^{(n)}, \ldots, \lambda_n^{(n)}\big)\,.$$

From the result on the density of states, we can anticipate that $\lambda_{\max}^{(n)}$ converges to 2 as $n \to \infty$, and this is indeed the case. The fluctuations of $\lambda_{\max}^{(n)}$ around

its limit 2 are interesting because they exhibit an anomalous scaling behavior (which is not the one corresponding to a central limit theorem) and the distribution of the fluctuations is original in the sense that they are not Gaussian but they follow the so-called Tracy-Widom distribution. More exactly, if $\lambda_{\max}^{(n)}$ is the maximal eigenvalue of $\mathbf{M}/\sqrt{n}$ where $\mathbf{M}$ is a GOE(n)-matrix (resp. a GUE(n)-matrix), then

$$n^{2/3}(\lambda_{\max}^{(n)} - 2)$$

converges in distribution as $n \to +\infty$ to a Tracy-Widom distribution of type 1 (resp. type 2). The type-1 Tracy-Widom distribution has the pdf p_{TW1} such that:

$$\int_{-\infty}^{y} p_{\mathrm{TW1}}(x)dx = \exp\Big(-\frac{1}{2}\int_{y}^{\infty} (\varphi(x) + (x-y)\varphi^2(x))\, dx \Big) ,$$

where φ is the solution of the Painlevé equation

$$\varphi''(x) = x\varphi(x) + 2\varphi(x)^3, \qquad \varphi(x) \overset{x\to+\infty}{\approx} \mathrm{Ai}(x) , \tag{6.11}$$

with Ai being the Airy function. The expectation is $\int x p_{\mathrm{TW1}}(x)dx \approx -1.21$ and the variance is approximately 1.61.
The type-2 Tracy-Widom distribution has the pdf $p_{\mathrm{TW2}}(x)$ such that

$$\int_{-\infty}^{y} p_{\mathrm{TW2}}(x)dx = \exp\Big(-\int_{y}^{\infty} (x-y)\varphi^2(x)dx \Big) .$$

The expectation is $\int x p_{\mathrm{TW2}}(x)dx \approx -1.77$ and the variance is approximately 0.81.

6.6.4 Wishart model

As a warm-up towards the statistical analysis of singular values of random matrices (with no special symmetry), we consider the Wishart model. Let $\mathbf{A}$ be a $m \times n$ real random matrix whose coefficients are independent and identically distributed with the Gaussian distribution with mean zero and variance one. Let us consider the symmetric $n \times n$ real random matrix

$$\mathbf{W} = \frac{1}{m}\mathbf{A}^T\mathbf{A} = \left(\frac{1}{m}\sum_{k=1}^{m} A_{ki}A_{kj} \right)_{i,j=1,\dots,n}$$

and let us denote by $(\lambda_1^{(n)}, \dots, \lambda_n^{(n)})$ the eigenvalues of $\mathbf{W}$. If $m/n \to c > 0$ as $n \to \infty$, then the density of states $\rho^{(n)}$ of $\mathbf{W}$,

$$\rho^{(n)}([a,b]) = \frac{1}{n}\mathrm{Card}\left\{ j = 1, \dots, n\,,\, \lambda_j^{(n)} \in [a,b] \right\} ,$$

converges to the measure

$$\rho(\lambda) = (1-c)^+\delta(\lambda) + \frac{1}{2\pi\lambda}\sqrt{[(\lambda-\lambda_-)(\lambda_+-\lambda)]^+}\,,$$

with $x^+ = \max(x,0)$, $\lambda_\pm = (1+c) \pm 2\sqrt{c}$. In the case $m = n$ we find

$$\rho(\lambda) = \frac{1}{2\pi\sqrt{\lambda}}\sqrt{4-\lambda}\mathbf{1}_{[0,4]}(\lambda)\,. \tag{6.12}$$

Taking into account that the eigenvalues of $\mathbf{W}$ are the squares of the singular values of $\mathbf{A}$, this result shows that the density of states of the singular values of $\mathbf{A}$ converges to the quarter-circle law with density $\rho_{\mathrm{qc}}(\sigma) = \frac{1}{\pi}\sqrt{4-\sigma^2}\mathbf{1}_{[0,2]}(\sigma)$.

References

1. M. Abramowitz and I. Stegun, *Handbook of mathematical functions*, Dover Publications, New York, 1965.
2. R. Adler, *The Geometry of Random Fields*, SIAM, Philadelphia, 2010.
3. R. Adler and J. Taylor, *Random Fields and Geometry*, Springer, New York, 2007.
4. H. Ammari, E. Bonnetier, and Y. Capdeboscq, *Enhanced resolution in structured media*, SIAM J. Appl. Math., 70 (2009), pp. 1428-1452.
5. H. Ammari, J. Garnier, and K. Sølna, *A statistical approach to target detection and localization in the presence of noise*, Waves in Random and Complex Media, 22 (2012), pp. 40-65.
6. A. P. Berens, *NDE reliability data analysis*, ASM Handbook, Vol. 17, 689-701, 1989.
7. G. Beylkin, M. Oristaglio, and D. Miller, *Spatial resolution of migration algorithms*, in 14th International Symposium on Acoustical Imaging, Proceedings, edited by A. J. Berkhout, J. Ridder, and L. F. van der Walls, pp. 155-167, Plenum, New York, 1985.
8. D. T. Blackstock, *Fundamentals of physical acoustics*, Wiley, New York, 2000.
9. N. Bleistein, J. K. Cohen, and J. W. Stockwell Jr, *Mathematics of multidimensional seismic imaging, migration, and inversion*, Springer Verlag, New York, 2001.
10. L. Borcea, T. Callaghan, J. Garnier, and G. Papanicolaou, *A universal filter for enhanced imaging with small arrays*, Inverse Problems 26 (2010), 015006.
11. M. Born and E. Wolf, *Principles of optics*, Cambridge University Press, Cambridge, 1999.
12. L. Breiman, *Probability*, Addison-Wesley, Reading, 1968; reprinted by Society for Industrial and Applied Mathematics, Philadelphia, 1992.
13. A. Chai, M. Moscoso, and G. Papanicolaou, *Imaging strong localized scatterers with sparsity promoting optimization*, SIAM J. Imaging Sciences, 7 (2014), pp. 1358-1387.
14. D. H. Chambers and J. G. Berryman, *Analysis of the time-reversal operator for a small spherical scatterer in an electromagnetic field*, IEEE Trans. Antennas and Propagation, 52 (2004), pp. 1729-1738.

15. D.H. Chambers and J.G. Berryman, *Time-reversal analysis for scatterer characterization*, Phys. Rev. Lett., 92 (2004), 023902.
16. J. Chen and G. Schuster, *Resolution limits of migrated images*, Geophysics, 64 (1999), pp. 1046-1053.
17. R. Courant and D. Hilbert, *Methods of mathematical physics*, Wiley, New York, 1991.
18. A. C. Davison, *Statistical Models*, Cambridge University Press, Cambridge, 2003.
19. A. J. Devaney, *A filtered backpropagation algorithm for diffraction tomography*, Ultrasonic Imaging, 4 (1982), pp. 336-350.
20. L. C. Evans, *Partial differential equations*, 2nd edition, American Mathematical Society, Providence, 2010.
21. M. Fink, *Time reversed acoustics*, Physics Today, 20 (1997), pp. 34-40.
22. J.-P. Fouque, J. Garnier, G. Papanicolaou, and K. Sølna, *Wave propagation and time reversal in randomly layered media*, Springer, New York, 2007.
23. J. Garnier and K. Sølna, *Coupled paraxial wave equations in random media in the white-noise regime*, Ann. Appl. Probab., 19 (2009), pp. 318-346.
24. J. Garnier and K. Sølna, *Scaling limits for wave pulse transmission and reflection operators*, Wave Motion, 46 (2009), pp. 122-143.
25. J. Garnier and K. Sølna, *Fourth-moment analysis for beam propagation in the white-noise paraxial regime*, Archive on Rational Mechanics and Analysis, 220 (2016), pp. 37-81.
26. I. I. Gihman and A. V. Skorohod, *The theory of stochastic processes*, Springer-Verlag, Berlin, 1974. vol. 1.
27. G. H. Golub and C. F. Van Loan, *Matrix computations (3rd ed.)*, Johns Hopkins, Baltimore, 1996.
28. C. Gomez, *Time-reversal superresolution in random waveguides*, SIAM Multiscale Model. Simul., 7 (2009), pp. 1348-1386.
29. I. M. Johnstone, *On the distribution of the largest eigenvalue in principal components analysis*, Ann. Statist., 29 (2001), pp. 295–327.
30. S. M. Kay, *Fundamentals of Statistical Signal Processing, Detection Theory*, Englewood Cliffs, Prentice-Hall, 1998.
31. G. Lerosey, J. de Rosny, A. Tourin, and M. Fink, *Focusing beyond the diffraction limit with far-field time reversal*, Science, 315 (2007), pp. 1120-1122.
32. P. D. Letourneau, *Fast algorithms and imaging in strongly scattering media*, PhD thesis, Stanford University, 2013.
33. V. A. Marcenko and L. A. Pastur, *Distributions of eigenvalues of some sets of random matrices*, Math. USSR-Sb., 1 (1967), pp. 507–536.
34. M. L. Mehta, *Random Matrices*, Academic Press, San Diego, 1991.
35. B. Perthame and L. Vega, *Energy concentration and Sommerfeld condition for Helmholtz equation with variable index at infinity*, Geom. Funct. Anal., 17 (2008), pp. 1685-1707.
36. J. Seberry, B. J. Wysocki, and T. A. Wysocki, *On some applications of Hadamard matrices*, Metrika, 62 (2005), pp. 221–239.
37. K. Wapenaar and J. Fokkema, *Green's function representations for seismic interferometry*, Geophysics, 71 (2006), pp. SI33-SI46.
38. R. Wong, *Asymptotic approximations of integrals*, SIAM, Philadelphia, 2001.

Index