§1 INTRODUCTION TO INVERSE PROBLEMS

inverse problems are defined (as term indicates) as inverse of forward problem (direct problem)

INPUT ⇒ SYSTEM ⇒ OUTPUT

CAUSE / SOURCE — EFFECT

direct / forward problem: we determine effect from cause

inverse problem: we determine cause from effect (we wish to compute information about hidden / unavailable data from outside (or otherwise accessible) measurements)

ill posedness of inverse problems

inverse problems belong to class of ill-posed problems

Jacques Hadamard (1932) defined a (linear) problem to be well-posed if it satisfies:

- existence: there exists at least one solution
- uniqueness: there must be only one solution
- stability: solution depends continously on data

forward problem is assumed to be well-posed

if inverse problem violates one or more of these requirements it is ill-posed example: existence condition

$$x = 1$$
$$2x = 11/5$$

$$\underset{x}{\text{minimize}} \left\| \begin{pmatrix} 1 \\ 2 \end{pmatrix} x - \begin{pmatrix} 1 \\ 11/5 \end{pmatrix} \right\|_2^2 = \underset{x}{\text{minimize}} \left((x-1)^2 + (2x - 11/5)^2 \right)$$

unique solution: $x = 54/50$

mathematically, existence of solution can be enforced by enlarging solution space

uniqueness condition usually more critical; if problem has more than one solution, then information is usually missing

typically resolved by reformulation of problem (additional prior knowledge about expected solution is built into model)

$$x_1 + x_2 = 1$$

has infinitely many solutions

add additional requirement that ℓ^2-norm of $x = (x_1, x_2)$ is minimum

unique solution: $x = (1/2, 1/2)$

stability condition more delicate: violation implies that small perturbations of data can produce arbitrarily large perturbations in solution; practically, solution is impossible to compute because any measurement or numerical computation is polluted by unavoidable errors

Lanczos (1986): "A lack of information cannot be remedied by mathematical trickery"

example: least squares problem

$$\text{minimize}_{x \in \mathbb{R}^2} \quad f(x) = \|Ax - y^\delta\|_2$$

$$A = \begin{pmatrix} 0.16 & 0.10 \\ 0.17 & 0.11 \\ 2.02 & 1.29 \end{pmatrix}, \qquad y^\delta = Ax^* + \eta$$

$$x^* = (1,1), \qquad \eta = (0.01, -0.03, 0.02)$$

$$x_{sol} = (A^TA)^{-1}A^Ty^\delta \approx (7.01, -8.40) \neq x^*$$

$$f(x_{sol}) \approx 0.02$$

observation:

$$A \begin{pmatrix} -1 \\ 1.57 \end{pmatrix} = \begin{pmatrix} -0.0030 \\ 0.0027 \\ 0.0053 \end{pmatrix} \approx 0$$

$\rightsquigarrow$ $z = \begin{pmatrix} -1 \\ 1.57 \end{pmatrix}$ is almost in null A

remedy: modify problem by enforcing upper bound $\varepsilon > 0$ on $\|x\|_2$

$$\text{minimize}_{x \in \mathbb{R}^n} \quad \|Ax - b^\delta\|_2$$

$$\text{subject to} \quad \|x\|_2 \leq \varepsilon$$

def (mathematical model): let $\mathcal{X}, \mathcal{Y}$ be normed spaces; a mathematical model is a (linear or nonlinear) mapping $\mathcal{K}: \mathcal{X} \to \mathcal{Y}$, from "causes" (parameters) $\mathcal{X}$ to "effects" $\mathcal{Y}$

direct problem: we compute effect $y \in \mathcal{Y}$ for a given cause $x \in \mathcal{X}$, i.e., we evaluate $\mathcal{K}x$ for $x \in \mathcal{X}$

inverse problem: we find cause $x \in \mathcal{X}$ for given effect $y \in \mathcal{Y}$ (data, observable), such that $\mathcal{K}x = y$

def (well-posedness): let $\mathcal{K}: \mathcal{X} \to \mathcal{Y}$ be a (linear or nonlinear) mapping between the normed spaces $\mathcal{X}$ and $\mathcal{Y}$; we call the problem $(\mathcal{K}, \mathcal{X}, \mathcal{Y})$ properly posed or well-posed if following properties hold:

- existence: for every $y \in \mathcal{Y}$ there is (at least) one $x \in \mathcal{X}$ such that $\mathcal{K}x = y$
- uniqueness: for every $y \in \mathcal{Y}$ there is at most one $x \in \mathcal{X}$ with $\mathcal{K}x = y$
- stability: solution x depends continously on y; that is, for every sequence $(x_n) \subset \mathcal{X}$ with $\mathcal{K}x_n \to \mathcal{K}x$ $(n \to \infty)$, it follows that $x_n \to x$ $(n \to \infty)$

is one of these definitions violated, we call problem $(\mathcal{K}, \mathcal{X}, \mathcal{Y})$ ill-posed example:

$$\Delta u(x_1, x_2) := \partial_{11} u(x_1, x_2) + \partial_{22} u(x_1, x_2) = 0 \text{ in } \mathbb{R} \times [0, \infty[$$

$$u(x_1, 0) = f(x_1), \quad \partial_2 u(x_1, 0) = g(x_1)$$

for $f(x_1) = 0$, $g(x_1) = \dfrac{\sin(nx_1)}{n}$, (unique) solution is given by

$$u(x_1, x_2) = K^{-1} g(x_1) = \frac{1}{n^2} \sin(nx_1) \sinh(nx_2)$$

we have

$$\sup_{x_1 \in \mathbb{R}} \{|f(x_1)| + |g(x_1)|\} = \frac{1}{n} \to 0 \quad \text{for} \quad n \to \infty$$

but

$$\sup_{x_1 \in \mathbb{R}} |u(x_1, x_2)| = \frac{1}{n^2} \sinh(nx_2) \to \infty \quad \text{for } n \to \infty$$

for all $x_2 > 0$

conclusion: error in data tends to zero, while error in solution tends to infinity

$\Rightarrow$ solution does not depend continuously on data

$\Rightarrow$ problem (K, X, Y) is improperly posed

example: differentiation

direct problem: find antiderivative g of a given continuous function f, i.e., compute

$$(*) \qquad K f(x) = \int_0^x f(t)\, dt = g(x), \quad x \in [0,1]$$

with g in $Y = \{\varphi \in C^1(0,1) : \varphi(0) = 0\}$

we have, $f(t) = d_t\, g(t)$

inverse problem: given a continously differentiable function g on $[0,1]$, determine f

integral operator $K: C(0,1) \to \mathcal{Y}$ is bijective and inverse $K^{-1}: \mathcal{Y} \to C(0,1)$ exists;
this implies that $\|K^{-1}\|_{\mathcal{Y}\to C(0,1)} \leq 1$ and by that, that problem $(K, C(0,1), \mathcal{Y})$ is well posed

now suppose we gained g through measurements

$$g^\delta = g + \eta$$

issues:
- $g^\delta \notin \mathcal{Y} = \{\varphi \in C^1(0,1) : \varphi(0) = 0\}$
- we cannot apply K^{-1} to g

we have to extend image of K to allow for perturbations

$$K: L^\infty(0,1) \to L^\infty(0,1)$$

we will see that $(K, L^\infty(0,1), L^\infty(0,1))$ is an ill-posed problem;

$$g^\delta(x) = g(x) + \delta \sin(nx) \quad , \quad \delta > 0$$

notice that $g(x) \in \{\varphi \in C^1(0,1) : \varphi(0) = 0\}$

let f^δ represent solution of (*)
we have $\|g^\delta - g\|_{L^\infty(0,1)} = \delta$

on other hand,

$$\|\mathfrak{g}^\delta - \mathfrak{g}\|_{L^\infty(0,1)} = \|(g^\delta)' - (g)'\|_{L^\infty(0,1)} = n\delta$$

for $\delta = \delta_n = n^{-1/2}$, we have

$$\|g^\delta - g\|_{L^\infty(0,1)} \xrightarrow{n\to\infty} 0 \qquad \|\mathfrak{g}^\delta - \mathfrak{g}\|_{L^\infty(0,1)} \xrightarrow{n\to\infty} \infty$$

central finite differences

$$\partial^h_\pm g(x) := \frac{1}{2h}\left(g(x+h) - g(x-h)\right)$$

suppose exact data g lives in $C^\ell(0,1)$, $\ell \geq 3$

$$g(x+h) = g(x) + hg'(x) + \frac{1}{2}h^2 g''(x) + \frac{1}{6}h^3 g'''(z^+)$$

$$g(x-h) = g(x) - hg'(x) + \frac{1}{2}h^2 g''(x) - \frac{1}{6}h^3 g'''(z^-)$$

with $z^+ \in [x, x+h]$, $z^- \in [x, x-h]$

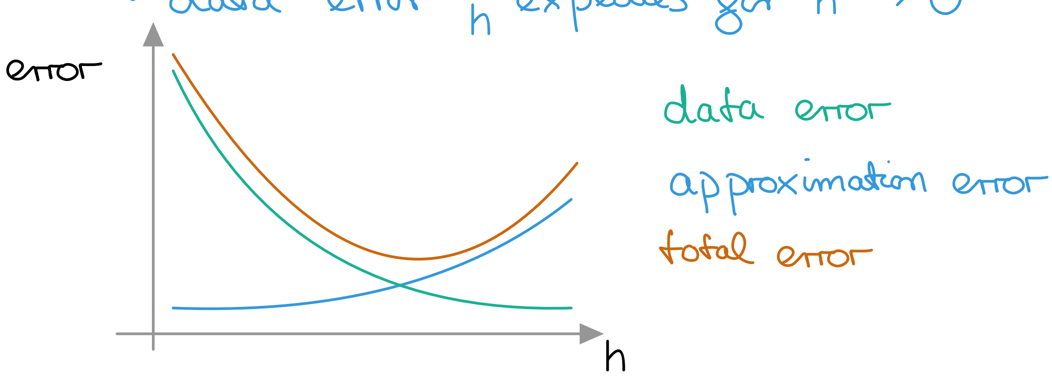
$$\|\mathfrak{g} - \partial^h_\pm g\|_{L^\infty} = \|g' - \partial^h_\pm g\|_{L^\infty} \leq \frac{h^2}{6}\|g'''\|_{L^\infty} = \frac{h^2}{6}\|\mathfrak{g}''\|_{L^\infty}$$

$$\|\mathfrak{g} - \partial^h_\pm g^\delta\|_{L^\infty} \leq \frac{h^2}{6}\|\mathfrak{g}''(z^+)\|_{L^\infty} + \frac{\bar{\delta}}{h}$$

since $\|\partial^h_\pm(g^\delta - g)\|_{L^\infty} \leq \frac{\bar{\delta}}{h}$, where $\|\eta\|_{L^\infty} \leq \bar{\delta}$

conclusion:

- approximation error $\frac{h^2}{6}\|\mathfrak{g}''\|_{L^\infty}$ decreases for $h \to 0$
- data error $\frac{\bar{\delta}}{h}$ explodes for $h \to 0$



examples for inverse problems

ⓘ linear regression and least squares

$$y^\delta = Kx + \eta, \qquad K \in \mathbb{R}^{m,n}$$

direct problem: "apply" K to x ("matvec")

inverse problem: given y^δ and K, find x

case (i) $n < m$ and $K: \mathbb{R}^n \to \operatorname{range} K \subsetneq \mathbb{R}^m$

due to noise perturbation $y^\delta \notin \operatorname{col} K = \operatorname{range} K$

we typically have no solution

(ii) $m < n$ and $K: \mathbb{R}^n \to \mathbb{R}^m$

there are several solutions

(iii) $m = n$, there exists $K^{-1}: \mathbb{R}^n \to \mathbb{R}^n$, and

$\kappa = \operatorname{cond} K = \frac{\lambda_1}{\lambda_n} \gg 1$

naive reconstruction: $x_{SOL} = K^{-1} y^\delta = x + K^{-1}\eta$

$\|K^{-1}\eta\|_2 \approx \|\eta\|_2 / \lambda_n$ can be arbitrarily large

find best approximation (deterministic setting)

$$\operatorname{minimize}_{x \in \mathbb{R}^n} \quad f(x) = \| y^\delta - Kx \|_2^2$$

optimality conditions: $2K^TKx^* - 2K^Ty = 0$

$$\tilde{x} = \arg\min_{x \in \mathbb{R}^n} f(x) = (K^TK)^{-1}K^Ty$$

we can compute unique and stable solution by regularization

$$\operatorname{minimize}_{x \in \mathbb{R}^n} \quad f(x) = \frac{1}{2}\|y^\delta - Kx\|_2^2 + \frac{\alpha}{2}\|Lx\|_2^2$$

statistical setting

question asked: not "what is the value of this variable" but "what is our information about this variable"

statistical method: does not produce only single estimates; rather, attempt to remove ill-posedness by restating inverse problem as a well-posed extension in a larger space of probability distributions

we rephrase inverse problem as a question of statistical inference

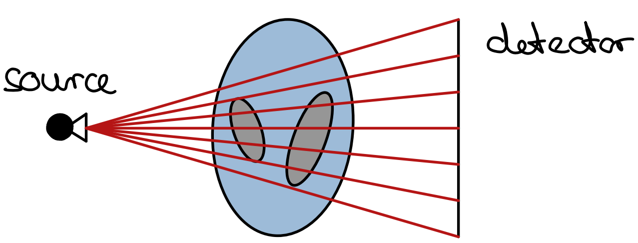
$$Y_{OBS} = KX + N \qquad (*)$$

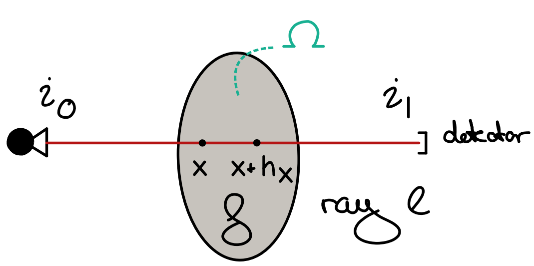
allows us to

- model noise through statistical properties
- encode a priori knowledge of unknown in form of a probability distribution that assigns higher probability to those values of X we expect to see

solution to $(*)$ is so-called <u>posterior distribution</u>, i.e., conditional probability distribution of X given measurements y_{OBS}

(ii) <u>computed tomography</u>

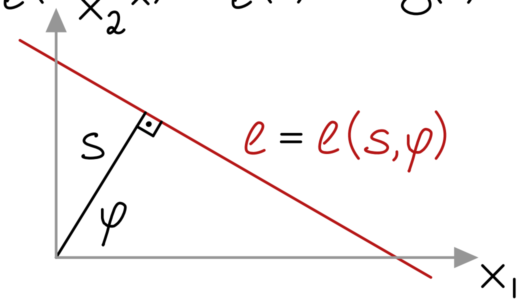




task: find $f(x)$, where $f(x) = 0$, if $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \notin \Omega$

$$i_\ell(x+h_x) - i_\ell(x) \propto \|h_x\| \, i_\ell(x)$$

$$\rightsquigarrow \quad i_\ell(x+h_x) - i_\ell(x) = -f(x) \, \|h_x\| \, i_\ell(x) \qquad (*)$$



$$\ell(s,\varphi) = \{ \gamma(t) : t \in \mathbb{R} \}$$

$$\gamma(t) = s\,\omega(\varphi) + t\,\omega^\perp(\varphi)$$

$$\omega(\varphi) = (\cos\varphi, \sin\varphi)^T \quad \text{and} \quad \omega^\perp(\varphi) = (-\sin\varphi, \cos\varphi)^T$$

$$x = \gamma(\tau), \qquad x + h_x = \gamma(\tau + h_\tau), \qquad \|h_x\| = h_\tau$$

$(*)$ becomes $$\frac{i_\ell(\gamma(\tau+h_\tau)) - i_\ell(\gamma(\tau))}{h_\tau} = -f(\gamma(\tau))\, i_\ell(\gamma(\tau))$$

$$\lim_{h_\tau \to 0} \frac{i_\ell(\gamma(\tau+h_\tau)) - i_\ell(\gamma(\tau))}{h_\tau} = d_t\, i_\ell(\gamma(t)) \Big|_{t=\tau}$$

$$d_t\, i_\ell(\gamma(t)) \Big|_{t=\tau} = -f(\gamma(\tau))\, i_\ell(\gamma(\tau))$$

$$\Leftrightarrow \quad f(\gamma(\tau)) = -\frac{1}{i_\ell(\gamma(\tau))}\, d_t\, i_\ell(\gamma(t)) \Big|_{t=\tau}$$

$$\Leftrightarrow \quad f(\gamma(\tau)) = -d_t \ln i_\ell(\gamma(t)) \Big|_{t=\tau}$$

$$\Leftrightarrow \quad \int_{\tau_S}^{\tau_D} f(\gamma(\tau))\, d\tau = -\ln i_1 + \ln i_0 = \ln \frac{i_0}{i_1}$$

mathematical formulation:

reconstruct f from line integrals across all rays that run through Ω

Radon transform:

$$\mathcal{R}f(s,\varphi) := \int_{\mathbb{R}} f(\underbrace{s\omega(\varphi) + t\omega^{\perp}(\varphi)}_{\gamma(t)})\,dt$$

measured data: $g(s,\varphi) = \ln \dfrac{i_0(s,\varphi)}{i_1(s,\varphi)}$

inverse problem: solve integral equation

$$\mathcal{R}f(s,\varphi) = g(s,\varphi)$$

for density f

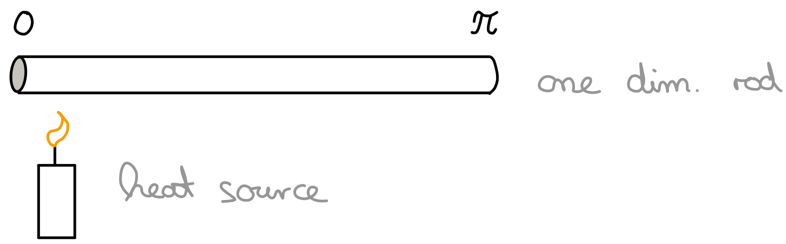
(iii) <u>inverse heat conduction problem</u>

3 variants

(i) we seek heat distribution at t_0, given heat distribution at $t_1 > t_0$

(ii) we seek heat distribution in interior of object based on heat flux across surface

(iii) we seek inhomogeneous thermal conductivity coefficients based on heat measurments on surface (problem is related to impedance tomography) problem (i)



$$\text{(HE)} \quad \begin{aligned} \partial_t u(t,x) &= \partial_{xx} u(t,x) \text{ in }]0,\pi[\times]0,1] \\ u(t,0) &= u(t,\pi) = 0, \qquad t \geq 0 \\ u(0,x) &= u_0(x), \qquad 0 \leq x \leq \pi \end{aligned}$$

direct problem: given $u(0,x) = u_0(x)$ at time $t=0$,
compute $u(t,x)$ for increasing time $t>0$

inverse problem: we seek $u_0(x) = \ (0,x)$ at time $t=0$,
given $u(1,x)$ at $t=1$

seperation of variables leads to (formal) solution

$$u(t,x) = \sum_{n=1}^{\infty} c_n \exp(-n^2 t) \sin(nx)$$

$$\text{with} \quad c_n = \frac{2}{\pi} \int_0^{\pi} u_0(y) \sin(ny)\, dy$$

interchanging summation and integration yields Fredholm integral equation of first kind

$$u(1,x) = \int_0^{\pi} K(x,y)\, u_0(y)\, dy \qquad (*)$$

with kernel $K(x,y) = \frac{2}{\pi} \sum_{n=1}^{\infty} \exp(-n^2) \sin(ny) \sin(nx)$

inverse problem: given $u(1,x)$, find u_0 from integral equation (*)

iv) diffusion in inhomogeneous medium

diffusion equation

describes medium

$$\partial_t u(t,x) = \nabla\cdot\sigma(x)\nabla u(t,x), \qquad x\in\Omega,\ t>0$$
$$u(0,x) = u_0(x)$$
$$\partial_n u(t,x) = 0$$

for stationary case

voltage

$$\nabla\cdot\sigma(x)\nabla u(x) = 0, \qquad x\in\Omega$$

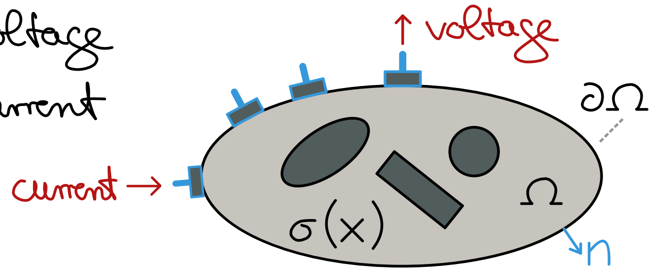
resistance

$$\sigma(x)\,\partial_n u = j(x), \qquad x\in\partial\Omega$$

current

u is voltage

j is current



"impedance tomography"

direct problem: solve boundary value problem for given boundary data $j(x)$ and parameters $\sigma(x)$

inverse problem: find $\sigma(x)$, given u and i on $\partial\Omega$
(parameter identification problem)

$$\Lambda_\sigma : j|_{\partial\Omega} \mapsto u|_{\partial\Omega}$$

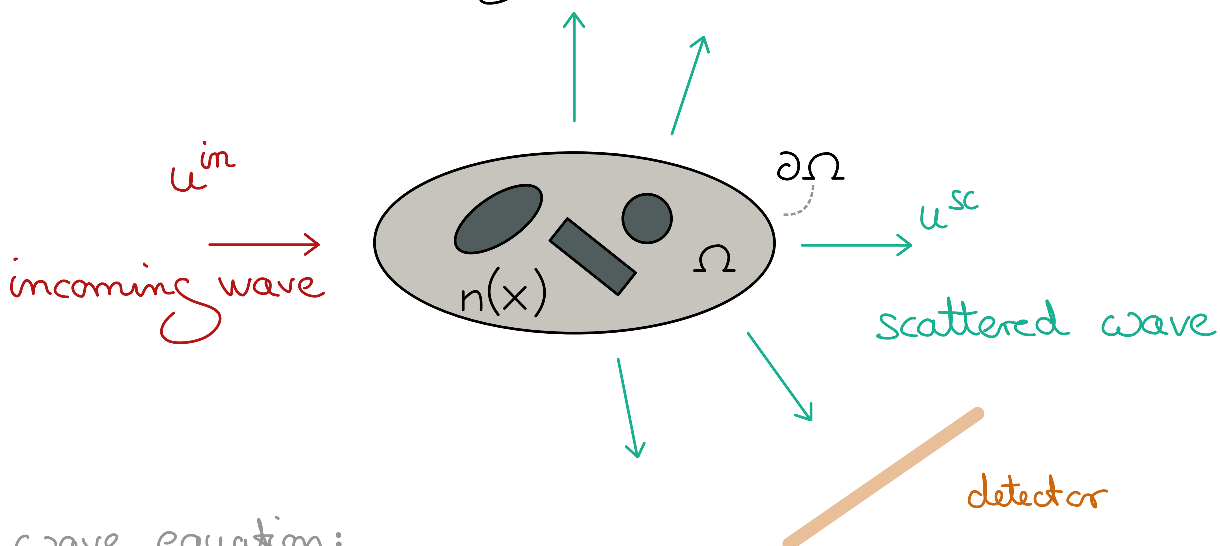
we define operator

$$\Phi : \sigma \mapsto \Lambda_\sigma$$

inverse problem: solve nonlinear equation

$$\Phi(\sigma) = \Lambda_{OBS}$$

(4) inverse scattering problem



wave equation:

$$\frac{1}{c^2(x)} \partial_{tt} u(t,x) = \Delta u(t,x), \qquad x \in \mathbb{R}^d, \quad t>0$$

wave speed

sound pressure

$$\Delta u(\omega,x) + k^2 n^2(x) u(\omega,x), \qquad (*)$$

where

$$u(\omega,x) := \left(\tfrac{2}{\pi}\right)^{-1/2} \int_{\mathbb{R}} u(t,x) \exp(-i\omega t)\, dt$$

with wave number $k = \omega / c_0$ and refractive index $n(x) = c_0 / c(x)$

superposition principle

incident wave

$$u(\omega,x) = u^{in}(\omega,x) + u^{sc}(\omega,x)$$

scattered wave

$$u^{in}(x) = u^{in}(r,x) = \exp(-ik r^T x)$$

$$\Delta u^{in}(r,x) + k^2 u^{in}(r,x) = 0 \qquad (*)$$

with $g(x) := 1 - n^2(x)$, $(*)$ becomes

$$\underline{\Delta u(r,x) + k^2 u(r,x)} = k^2 g(x) u(r,x)$$

with superposition principle and (*) we have

$$\Delta u^{sc}(r,x) + k^2 u^{sc}(r,x) = k^2 f(x)\, u(r,x) \qquad (*)$$

using fundamental solution (Green's function G) of Helmholtz equation yields

$$u^{sc}(r,x) = k^2 \int_\Omega G(x-y)\, f(y)\, u(r,y)\, dy \qquad \text{(LSI)}$$

with $G(x) = -\frac{1}{4} i\, \mathcal{H}(k\|x\|)$ for $d=2$, $x \neq 0$

Hankle function of first kind

$$G(x) = -\frac{1}{4\pi} \frac{\exp(ik\|x\|)}{\|x\|} \quad \text{for } d=3, \quad x \neq 0$$

direct problem: compute u^{sc}, given f and u

inverse problem: given $u^{sc}(r,x) = u(r,x) - u^{in}(r,x)$ on detector (i.e., for all $x \in D$) for all directions r find f, and by that n

Born – Rytov approximation

$$u^{sc}(r,x) = k^2 \int_\Omega G(x-y)\, f(y) \exp(-ik r^T y)\, dy$$