This is a summary of the advances made, during the period of June 1994 to September 1995, in the investigation of the inverse scattering problem for the Helmholtz equation in two dimensions. The principal results presented here are two stable methods for the solution of the fully nonlinear problem. The underlying physics employed is the so-called uncertainty principle: it is increasingly difficult to determine features in the scatterer as their sizes become decreasingly smaller than a half of a wavelength.

This new approach belongs to a more general principle: some of the equations in an ill-posed nonlinear system are essentially linear. These equations may be first solved to produce a least-squares solution, and the approximate solution may further be used to linearize other remaining equations with stronger nonlinearity.

Recursive Linearization for Inverse Scattering

Yu Chen
Research Report YALEU/DCS/RR-1088
October 3, 1995

The author was supported in part by the ONR under the grants ONR N00014-93-1-0114 and ONR N00014-89-J-1527.

Keywords: *Inverse Scattering, Helmholtz Equation, Riccati Equation, Scattering Matrix, Recursive Linearization, Uncertainty Principle, Ill-posed Problems, Nonlinear Equations.*

1 Introduction

The solution of an inverse scattering problem requires, in essence, an inversion of a nonlinear mapping. There are two major difficulties associated with this nonlinear problem: ill-posedness and local minima, neither of which has been addressed satisfactorily. It turns out that the ill-posedness of the inverse problem can be beneficially used to solve it. It means that, due to ill-posedness of the problem, not all equations in the nonlinear system are strongly nonlinear, and that when solved in a proper order, they can be recursively reduced to a collection of linear problems.

The purpose of this report is to summarize the advances made, during the period of June 1994 to September 1995, in the investigation of the inverse scattering problem for the Helmholtz equation in two dimensions. The principal results presented here are two stable methods that solve the fully nonlinear acoustic inverse scattering problem.

The plan of the paper is as follows: in Section 2 we summerize several fundamental principles for forward and inverse scattering. In Section 3, we describe the inversion algorithms. Finally in Section 4, we extend the result to other inverse problems and more general ill-posed nonlinear problems. The numerical implementations and results of the algorithms will be presented in subsequent papers.

2 The Analytical Apparatus

In this section, we summarize several fundamental principles for the forward and inverse scattering problems for the Helmholtz equation in two dimensions

$$\Delta\phi(x) + k^2(1 + q(x))\phi(x) = 0. \tag{1}$$

2.1 The Scattering Problem

In (1), we assume that k is a positive number, and q is a smooth function with compact support $\Omega \subset R^2$; we will be referring to the function q as the scatterer, or the forward model. We will be considering solutions of (1) of the form

$$\phi(x) = \phi_0(x) + \psi(x), \tag{2}$$

with ϕ_0 the incident, or the incoming, field satisfying in Ω the homogeneous Helmholtz equation

$$\Delta\phi_0(x) + k^2\phi_0(x) = 0, \tag{3}$$

and with ψ the scattered field subject to the outgoing (Sommerfeld) radiation condition

$$\lim_{r\to\infty} \sqrt{r}\left(\frac{\partial\psi}{\partial r} - ik\psi\right) = 0. \tag{4}$$

We will be referring to the determination of the scattered field from a given incoming field as the forward scattering problem. It is well-known (see, for example, [4]) that the forward scattering problem is well-posed. Moreover, the problem can be reformulated as the so-called Lippmann-Schwinger equation for the scattered field ψ

$$\psi(x) = -k^2 \int_\Omega G_k(x,\xi)q(\xi)(\phi_0(\xi) + \psi(\xi))d\xi, \tag{5}$$

or, to explicitly express the dependence of the fields on k,

$$\psi(x,k) = -k^2 \int_\Omega G_k(x,\xi)q(\xi)(\phi_0(\xi,k) + \psi(\xi,k))d\xi. \tag{6}$$

Losely speaking, the inverse scattering problem is to determine the scatterer q inside the domain Ω from measurements of the scattered field outside Ω. The inverse problem is nonlinear since both the scatterer and the scattered field is unknown inside Ω, and there is a product of them in the integral equation (5).

2.2 Three Special Cases

There are three special cases where the nonlinear relationship between the scatterer and the scattered field is essentially linear. In other words, the nonlinear inverse problem can be linearized when

1. q is small, or
2. Ω is small, or
3. k is small.

In each case, the maximum norm of the linear mapping $G_{(k,\Omega,q)} : C(\Omega) \mapsto C(R^2)$, defined by the formula

$$(G_{(k,\Omega,q)} \cdot \psi)(x) = k^2 \int_\Omega G_k(x,\xi)q(\xi)\psi(\xi)d\xi, \tag{7}$$

is small. Consequently, the scattered field is weak (see (5)), and the Born approximation

$$\psi(x) = -k^2 \int_\Omega G_k(x,\xi)q(\xi)\phi_0(\xi)d\xi \tag{8}$$

linearizes the relationship between q and ψ. It turns out that a simple understanding and a proper use of the three special cases are all that is required to construct a stable algorithm for the solution of the fully nonlinear problem. Further analyses of the linear operator $G_{(k,\Omega,q)}$, or equivalently, of the behavior of forward scattering, will be made in Section 2.7 which enable us to design yet another algorithm for both inverse scattering and impedance tomography (see [1], [2], [3]).

Remark 2.1 *The maximu norm of* $G_{(k,\Omega,q)}$ *is bounded, for example, by the number*

$$\delta = \frac{\frac{5}{2} + \left|\ln\left(\mid k \mid \cdot \sqrt{\frac{\mu}{\pi}}\right)\right|}{4} \mid k \mid^2 \cdot \|q\|_\infty \cdot \mu; \tag{9}$$

see Lemmas 2.9 and 2.18 in [5]. However, it is technically advantageous here to state three separate conditions on which the norm is small.

2.3 The First Special Case

In this section, we restate some well-known results regarding to the special case of small q where the relationship between q and ψ is linear.

At a fixed frequency $k > 0$, the most complete set of acoustic measurements can be obviously obtained by the acquisition of ψ on a circle (or a closed curve) containing Ω, for every linearly independent incoming field. An example of such a set of incoming fields is the plane waves of the form

$$\phi_0(x, y; \beta) = e^{ik(x\cos\beta + y\sin\beta)} \tag{10}$$

for every $\beta \in [0, 2\pi]$. Denoting by $\psi(r, \theta; \beta)$ the scattered field corresponding to the incoming field (10), we are led to the following definitions.

Definition 2.2 *An aperture (of acoustic measurement) is an area in the square* $[0, 2\pi] \times [0, 2\pi]$*; the full aperture is the entire square. A Fourier aperture is an area in the (two-dimensional) Fourier space; a function* $g \in L^2(\Omega)$ *is said to be specified on a Fourier aperture* E *if its Fourier transform* $\hat{g}(m, n)$ *is known for all modes* $(m, n) \in E$.

Assuming that the scatterer q is small, and denoting by $\psi_\infty(\theta, \beta)$ the standard far-field of the scattered field $\psi(r, \theta; \beta)$ obtained via the Born approximation (8), we can easily show that

$$\psi_\infty(\beta, \theta) = \frac{1}{2\pi} \int_\Omega q(x', y') e^{ik\{x'(\cos\beta - \cos\theta) + y'(\sin\beta - \sin\theta)\}} dx' dy' \tag{11}$$

In other words, for a small q, the far-field measurement $\psi_\infty(\beta, \theta)$ is the Fourier transform $\hat{q}(m, n)$ with

$$m = k(\cos\beta - \cos\theta), \tag{12}$$
$$n = k(\sin\beta - \sin\theta). \tag{13}$$

Therefore, the full-aperture far-field measurements

$$\{\ \psi_\infty(\beta,\theta),\ \text{ for all }\ (\beta,\theta)\in[0,2\pi]\times[0,2\pi]\ \} \tag{14}$$

are the Fourier transform $\hat{q}(m,n)$ for those modes (m,n) filling D_{2k}, the disk of radius $2k$ centered at the origin – a Fourier aperture of radius $2k$.

Remark 2.3 *Due to (12) and (13), the far-field $\psi_\infty(\beta,\theta)$ can be regarded as a function of (m,n). We therefore define the L^2 norm of the far-field $\psi_\infty(m,n)$ in a domain $E\subset D_{2k}$ by the formula*

$$\|\psi_\infty\|_2=\left[\int_E|\psi_\infty(m,n)|^2\cdot dm\cdot dn\right]^{\frac{1}{2}}. \tag{15}$$

2.4 The Uncertainty Principle

In this section, we extend results for the special case of small q to the general case where q is not small, and the relationship between q and ψ is nonlinear.

We know from the preceding section that, for small q, the full-aperture far-field measurements are the Fourier modes of q in the aperture D_{2k}. Therefore, in the inverse scattering, the scatterer q can be determined from the far-field measurements, for example, by a backward Fourier transform, with a resolution

$$\sigma=\frac{2\pi}{\text{radius of Fourier aperture}}=\frac{2\pi}{2k}=\frac{1}{2}\lambda, \tag{16}$$

where

$$\lambda=\frac{2\pi}{k} \tag{17}$$

is the wavelength.

Lemma 2.4 *(Uncertainty Principle: small q, far-field) Suppose that q is small. Then from the far-field measurements, we can not determine features of the scatterer that are smaller than a half of a wavelength.*

Obviously, the features smaller than a half of a wavelength correspond to Fourier modes $\hat{q}(m,n)$ with $\sqrt{m^2+n^2}>2k$. These higher-frequency components of q, in turn, correspond to evanescent modes, or non-propagating modes, in the scattered field, all of which decay exponentially while traveling to infinity. The following lemma is a reformulation of the preceding one.

Lemma 2.5 *Suppose that the scatterer q is small. Suppose further that $\delta\psi$ is the perturbation in the far-field due to a perturbation δq to the scatterer. Finally, suppose that the perturbed scatterer*

$$\tilde{q}=q+\delta q \tag{18}$$

is still small. Then

$$\|\delta\psi\|_2 = \|\delta q\|_2 \tag{19}$$

if the support of $\hat{\delta q}$ *is inside* D_{2k}*, and*

$$\|\delta\psi\|_2 = 0 \tag{20}$$

if the support of $\hat{\delta q}$ *is outside* D_{2k}*.*

It is not difficult to prove that full-aperture near-field measurements fill a Fourier aperture, which we denote by D_{2k}^+, slightly greater than the disk D_{2k} (the far-field Fourier aperture). In fact, the near-field aperture D_{2k}^+ is entirely embedded in a disk of radius $2k + O(k^{\frac{1}{3}})$.

More precisely, when the scatterer is perturbed outside the Fourier aperture D_{2k}^+, the perturbation in the near-field measurements is virtually zero (compare this with (20)); namely, these small features of the scatterer are non-observable. When the scatterer is perturbed in the transition zone – inside D_{2k}^+ but outside D_{2k}, the resulting perturbation in the measurements is weak compared to perturbation of the scatterer. Finally, when the scatter is perturbed inside D_{2k}, $\delta\psi$ is comparable to δq, as in the case of (19).

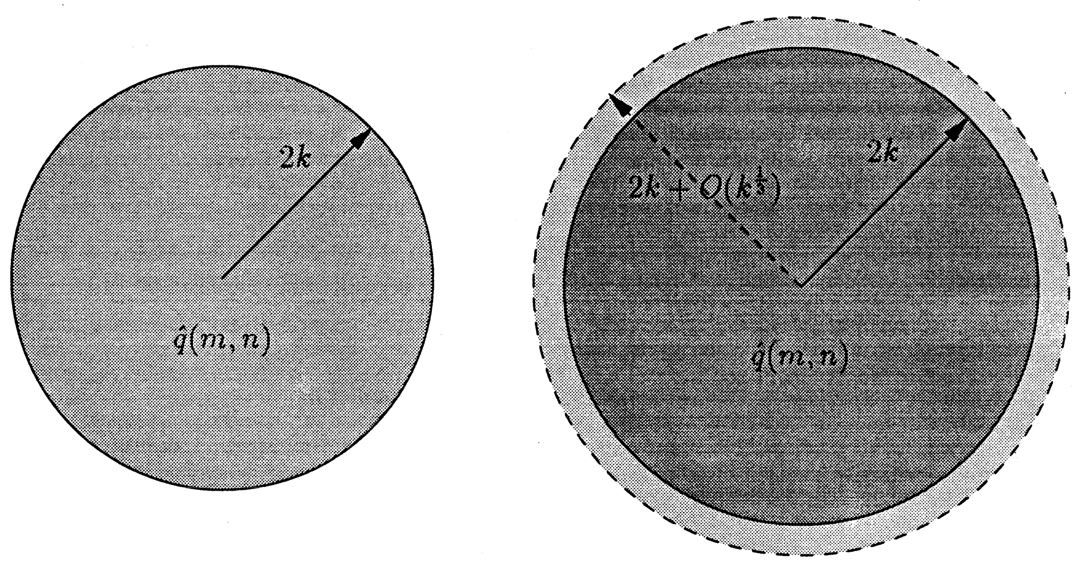

Figure 1. The Fourier apertures D_{2k} and D_{2k}^+

Remark 2.6 *In a physically meaningful inversion, it therefore makes no sense to recover from measurements the Fourier modes of the scatterer that are outside* D_{2k}^+*. It is variably difficult to recover the modes in the transition zone: it is easier to recover modes which are closer to the boundary of* D_{2k}*.*

We can prove, for example, using the Riccati equation for the scattering matrix developed in [5], that the above statements remain true for the general case where q is not small, and the relationship between q and ψ is nonlinear.

Lemma 2.7 *(Uncertainty Principle: near field, general case) It is increasingly difficult to determine features in the scatterer as their sizes become decreasingly smaller than a half of a wavelength.*

Remark 2.8 *Obviously, the Uncertainty Principle can be regarded as an equivalent formulation of the ill-posedness of the inverse scattering problem: the non-observable, small features of the scatterer belong to the null space of the nonlinear mapping which defines the inverse problem. In view of the uniqueness results – the scatterer can be uniquely determined by full-aperture measurements at essentially a single frequency k (see, for example, [4]), it would be reasonable to conclude that the scatterer can be stably determined in the Fourier aperture D_{2k}^{+}.*

2.5 Reformulating Scattering Problem

Denote by q_k the low-frequency part of q, corresponding to Fourier aperture D_{2k}^{+}, so that

$$\hat{q}_k(m,n) = \begin{cases} \hat{q}(m,n), & (m,n) \in D_{2k}^{+}, \\ 0, & (m,n) \notin D_{2k}^{+}. \end{cases} \tag{21}$$

The goal of inversion, in the lights of Remarks 2.6 and 2.8, is to stably obtain q_k within a reasonable precision. Since q_k is the part of q that is observable in the full-aperture near-field measurements, the original forward scattering model q can be replaced by q_k without essentially changing the measurements. In fact, we can show that at a given frequency k and for a given precision, the original scatterer q can be substituted by a smooth version of it, which produces the same scattered field as the original, to the prescribed precision. We therefore reformulate the original scattering problem (6) by an approximate one:

$$\psi(x,k) = -k^2 \int_\Omega G_k(x,\xi) q_k(\xi)(\phi_0(\xi,k) + \psi(\xi,k)) d\xi. \tag{22}$$

Definition 2.9 *To a scattering experiment at frequency k, a scatterer $\check{q}$ is said to look (essentially) the same as a scatterer $\check{q}$ if they produce essentially the same scattering measurements in the experiment.*

Definition 2.10 *A forward model $\tilde{q}$ is said to be observable, or an observable part of q, to a scattering experiment at frequency k if its L^2 norm is the smallest among the forward scattering models that look the same as the original q to the scattering experiment.*

Remark 2.11 *At frequency k, q_k looks the same as the original q to a full-aperture experiment, it is also observable to the full-aperture experiment. On the other hand, in an experiment of limited aperture, q_k may not be the observable forward model, but it looks the same as the observable.*

2.6 Continuity of q_k on k

Since q_k is the observable part of q at frequency k, it would be intuitive to think that, as the frequency is slightly changed, the change in the observable part of q should be small. In other words, q_k depends on k continuously. This is certainly the case when q is small. There the observable part of q, through far-field measurements, corresponds to the Fourier modes of q in aperture D_{2k}. Therefore, new Fourier modes added to $q_{k+\delta k}$ are those $\hat{q}(m,n)$ in the annulus

$$A(k,\delta k) = \{ \ (m,n), \ k \leq \sqrt{m^2+n^2} \leq k+\delta k \ \}. \tag{23}$$

Consequently, the perturbation in q_k, due to that in k, is small:

$$\|q_{k+\delta k} - q_k\|_2 = \|\widehat{q_{k+\delta k}} - \hat{q_k}\|_2 = \int_{A(k,\delta k)} \hat{q}(m,n)dmdn = O(\delta k). \tag{24}$$

We can show that q_k depends on k continuously, for full-aperture near-field measurements.

Lemma 2.12 *To a full-aperture, near-field experiment, the observable scattering model q_k depends continuously on k in the L^2 norm.*

We wish to carry this point further to the case of limited aperture. Denoting by $q_{k,l}$ the observable part of q corresponding to an experiment of a limited aperture, we remark that generally $q_{k,l}$ is not the same as q_k, and therefore, due to Definition 2.10,

$$\|q_{k,l}\|_2 < \|q_k\|_2. \tag{25}$$

Based on our experience with several special cases of limited aperture, we now make an assumption on the dependence of $q_{k,l}$ on k: we assume that Lemma 2.12 is also valid for scattering experiments with limited aperture, which we have observed in numerical experiments, and which can be proved in the special case of small q.

Observation 2.13 *$q_{k,l}$ depends continuously on k in the L^2 norm.*

2.7 Partial Linearization

The apparatus introduced in the preceding subsections will enable us to design a stable algorithm for the inverse problem. All these tools are actually built on the three special cases (see Section 2.2) where the nonlinear problem of inverse scattering can be linearized. There are other special cases where the scattered field can be made weak and therefore the relationship between the scatterer and the scattered field is essentially linear. There are yet more special linearities when the scattered field is not weak.

These additional special cases will allow us to speed up the stable inversion method, and to build other stable algorithms for other inverse problems, as well as for this inverse scattering problem for the Helmholtz equation. In this section, we analyze some of these special cases.

Let us consider the general case where q is not small. The following trivial result is formulated as a lemma for future reference.

Lemma 2.14 *A small perturbation δq in the scatterer results in a small perturbation $\delta\psi$ in the scattered field. Furthermore, up to the second order of the perturbations, $\delta\psi$ depends linearly on δq.*

2.7.1 Small Perturbation in Scattered Field

In general, if δq is not small, then $\delta\psi$ will not be small, and will depend on δq nonlinearly. But there are cases where $\delta\psi$ is small, and depends on δq, which is not assumed small, essentially linearly.

Lemma 2.15 *Suppose that q is the forward model in the Lippmann-Schwinger equation (5). Suppose further that q is perturbed by the amount δq which is not small. Finally, suppose that the perturbation in the scattered field ψ is small. Then,*

$$k^2 \int_\Omega G_k(x,\xi)\phi(\xi)\delta q(\xi)d\xi = O(\delta\psi). \tag{26}$$

Remark 2.16 *The lemma, whose proof is straightforward, implies that the perturbation in q lies essentially in the null space of the linear operator $G_\phi : C(\Omega) \mapsto C(R^2)$, defined by the formula*

$$(G_\phi \cdot \delta q)(x) = k^2 \int_\Omega G_k(x,\xi)\phi(\xi)\delta q(\xi)d\xi; \tag{27}$$

or, in the terms of linear algebra, δq is essentially orthogonal to the "rows" of G_ϕ, when δq is hardly observable. In the limiting case of δq being not at all observable, namely, when $\delta\psi = 0$, the orthogonality will be exact.

We have observed in numerical experiments that under the conditions of Lemma 2.15 and up to second order of $\delta\psi$, the perturbation of the scattered field is related to δq by a linear mapping; this can be proved for some special cases.

Observation 2.17 *Suppose that q is the forward model in Lippmann-Schwinger equation (5). Suppose further that q is perturbed by the amount δq which is not small. Finally, suppose that the perturbation $\delta\psi$ is small. Then, up to second order of the smallness, $\delta\psi$ depends on δq linearly:*

$$\delta\psi(x) = -k^2 \int_\Omega G_k(x,\xi) \cdot (\phi(\xi)\delta q(\xi) + q\delta\psi(\xi)) \cdot d\xi + O(\|\delta\psi\|_2^2). \tag{28}$$

Remark 2.18 *The Fourier modes of q outside the aperture D_{2k}^+ is not observable, and therefore belong to the null space of the linear operator G_ϕ. The Fourier modes of q in the transition zone – inside D_{2k}^+ but outside D_{2k} – are variably observable: $\delta\psi$ is weak compared to δq, and is weaker if the Fourier modes of q is perturbed near the boundary of D_{2k}^+ than near the boundary of D_{2k}. According to Observation 2.17, perturbations of q in the transition zone induce variable linearity between $\delta\psi$ and δq, with the latter not being assumed small.*

2.7.2 The Skin Effect

There are things we can do to the incoming field ϕ_0 in order to linearize the relationship between q and ψ. One of them is due to the fact that if the solution of the homogeneous Helmholtz equation (3) varies in a direction at a frequency higher than k, it decays or grows exponentially in perpendicular directions. Jaques Hadamard is probably the first to realize the potential hazards associated with such a phenomenon.

Let us consider a scatterer located in the unit disk for simplicity. If we have an incoming field, produced by sources outside the scatterer, which oscillates along the boundary at a frequency $m > k$, it will decay exponentially in the radial direction as it travels into the scatterer. In fact, such an incoming field has the form

$$\phi_0(r,\theta) = J_m(kr) \cdot e^{im\theta}. \tag{29}$$

For m substantially greater than k, the Bessel function $J_m(kr)$ decays fast as the radius r goes to zero

$$J_m(kr) \sim \frac{1}{\sqrt{2\pi m}} \left(\frac{e \cdot kr}{2m}\right)^m. \tag{30}$$

Therefore, the incoming field can only penetrate the skin of the scatterer which interacts with the incoming field, and produces a weak scattered field compared to the incoming. Thus, the Born approximation can be used to linearize the relationship between q and ψ.

As $m > k$ is closer to k, the skin gets thicker, and a greater part of the scatterer becomes observable, and the relationship between q and ψ turns more nonlinear.

2.7.3 The Gaussian Beam

Another type of the incoming field ϕ_0 we choose to linearize the relationship between q and ψ is the so-called Gaussian beam. We are not interested in describing what a Gaussian beam is, except that it is a solution of the homogeneous Helmholtz equation (3), which is essentially zero (dark) outside a sector of the form

$$\{ \ (r,\theta) \mid r \geq r_0, \ |\theta - \theta_0| \leq \frac{b}{k} \ \}, \tag{31}$$

with r_0, θ_0, b some constants, and with the origin of the beam a point in R^2. Therefore, the support of such an incoming field ϕ_0 looks like a beam from a search light over whose position and direction we have control.

Now suppose that the scatterer q is smooth, and that k is large. When the beam ϕ_0 is turn to the rim of the scatterer, only a controlled small portion of the scatterer will be illuminated. Consequently, the scattered field ψ will be weak, and again the Born approximation can be applied to linearize the relationship between the scatterer and the scattered field.

Obviously, other types of incoming fields, than the exact Gaussian beam, can also be used to illuminate a controlled portion of the scatterer.

3 Recursive Linearization for Inverse Scattering

We present here two inversion methods, one using multi-frequency, the other a single frequency, for the inverse scattering problem of the Helmholtz equation.

3.1 Recursive Linearization via Uncertainty Principle

The first breakthrough in solving the fully nonlinear inverse scattering problem occurred when we applied the Uncertainty Principle (see Section 2.4) to recursively linearize the nonlinear system (see Section 2.5)

$$\psi(x,k) = -k^2 \int_\Omega G_k(x,\xi) q_k(\xi)(\phi_0(\xi,k) + \psi(\xi,k)) d\xi. \tag{32}$$

In this section, we briefly describe the linearization procedures.

Suppose that the full-aperture near-field measurements are available at all frequencies $k > 0$ (full-spectrum). We discretize the k-space into nodes

$$k_1, \ k_2, \ k_3, \ \ldots, \tag{33}$$

with the first frequency k_1 so low that the relationship between q and ψ is essentially linear (see Section 2.2). Therefore, the linearized equation

$$\psi(x, k_1) = -k_1^2 \int_\Omega G_{k_1}(x, \xi)\phi_0(\xi, k_1) q_{k_1}(\xi) d\xi. \tag{34}$$

can be solved for q_{k_1}, the observable part of the scatterer. Since the measurements $\psi(x, k_1)$ are full-aperture, we feel that q_{k_1} should be stably determined (see Remark 2.8). This has been numerically confirmed.

Remark 3.1 *In numerical experiments, it is found that the lowest frequency k_1 can be chosen consistently so large that its corresponding wavelength is size of the scatterer.*

Now, by induction, suppose that we have recovered q_k, the observable part of the scatterer at frequency k, we wish to recover $q_{\tilde{k}}$ at a slightly greater $\tilde{k}$, or equivalently, the perturbation

$$\delta q = q_{\tilde{k}} - q_k. \tag{35}$$

Remark 3.2 *By definition of q_k (see Section 2.5), δq consists of Fourier modes of q in $D^+_{2\tilde{k}} \setminus D^+_{2k}$.*

This can certainly be achieved by employing standard perturbation analysis of the equation (32) on the parameter k. To this end, let us solve at the frequency $\tilde{k}$ the forward scattering problem

$$\check{\psi}(x, \tilde{k}) = -\tilde{k}^2 \int_\Omega G_{\tilde{k}}(x, \xi) \cdot q_k(\xi) \cdot (\phi_0(\xi, \tilde{k}) + \check{\psi}(\xi, \tilde{k})) \cdot d\xi, \tag{36}$$

with the forward model q_k. Since k is close to $\tilde{k}$, the perturbation δq is small; according to Lemma 2.14, the perturbation in scattered field

$$\delta\psi(x) = \psi(x, \tilde{k}) - \check{\psi}(x, \tilde{k}) \tag{37}$$

is also small (for x inside and outside Ω), and depends on δq linearly up to second order of the smallness. In other words, subtracting equation (36) from the equation

$$\psi(x, \tilde{k}) = -\tilde{k}^2 \int_\Omega G_{\tilde{k}}(x, \xi) \cdot q_{\tilde{k}}(\xi) \cdot (\phi_0(\xi, \tilde{k}) + \psi(\xi, \tilde{k})) \cdot d\xi, \tag{38}$$

we have, up to the second order of the smallness, Lippmann-Schwinger type of integral equation

$$\delta\psi(x) = -\tilde{k}^2 \int_\Omega G_{\tilde{k}}(x,\xi) \cdot [(\phi_0(\xi,\tilde{k}) + \check{\psi}(\xi,\tilde{k}))\delta q(\xi) + q_k \delta\psi(\xi)]d\xi. \tag{39}$$

For given q_k, $\check{\psi}$, which are indeed known, the integral equation can be solved to obtain a linear expression of $\delta\psi(x)$, at measurement points x, via δq. Denoting this linear procedure by $\mathcal{L}$, we have

$$\mathcal{L}(\delta q) = \delta\psi(x). \tag{40}$$

The linear problem can be solved, for the given right hand side $\delta\psi(x)$, to yield δq, and $q_{\tilde{k}}$ can be obtained via (35).

Remark 3.3 *Note that the linear operator $\mathcal{L}$, as well as $\delta\psi$, depends on the incoming field. In order to use full-aperture scattering data $\delta\psi$, linear equations (40) corresponding to all incoming fields should be solved simultaneously for δq.*

Remark 3.4 *Since all forward models that look the same (see Definition 2.9) as q_k satisfies the equation (32), the solution to the linear system (40) is obviously non-unique, and we need to solve a linear least-squares problem, which we know so well how to. In practice, the least-squares problem is solved with relative low precision, making an error in δq, whose Fourier components lie largely in the transition zone of the aperture $D^+_{2\tilde{k}}$. The error will be dealt with in subsequent steps of this recursive procedure. Eventually, the error, which corresponds to Fourier modes difficult to be observable at the present k, will be well within an aperture of a greater k, in which it will be well observable.*

Remark 3.5 *In our numerical implementation, equations (34), (40) were solved only once at each frequency, making an second order error there. The recovered q_k therefore should be close to, but will not be the original q_k which is the observable part of q at frequency k. Consequently, in going up to the next higher frequency $\tilde{k}$, added to δq will be not only Fourier modes in $D^+_{2\tilde{k}} \setminus D^+_{2k}$ (see Remark 3.2), but also those inside D^+_{2k} due to inexact calculation of q_k.*

Remark 3.6 *For inversion using near-field measurements, the step size in frequency k can be quite large, because the perturbation δq, though no longer small now, is small in the aperture D^+_{2k}, and is large only in the transition zone of $D^+_{2\tilde{k}}$ (see the preceding remark). Consequently, owing to Remark 2.18, $\delta\psi$ depends essentially linearly on this large component in δq, provided that the step size in k is not too large: the aperture D^+_{2k} should cover part of the transition zone of the next aperture $D^+_{2\tilde{k}}$.*

Numerical experiments further show that frequently k_j in (33) can be chosen such that the size of the scatterer is about j wavelengths. For instance, we may set $k_j = j$ for a scatterer inside a disk of diameter 2π.

The method of recursive linearization has been implemented, and the numerical results will be presented on a later date.

3.2 Layer Stripping via Skin Effect

The skin effect described in Section 2.7.2 can also be used to recursively linearize the nonlinear inverse problem. The advantages of such a procedure are largely that it uses a single frequency, and the computational effort may be reduced; although the procedure to be described here costs essentially the same as that of the multifrequency (see the preceding subsection) – they both require $O(N^5)$ operations for an $N \times N$ wavelengths problem in two dimensions.

The layer-stripping procedure is simple: in stead of having recursion in ascending direction of frequency k as in the method presented in the preceding subsection, the recursive linearization will be executed in the descending direction of the so-called propagation number m (see (29)).

Let us again consider a scatterer located in the unit disk for simplicity, and use the incoming fields of the form

$$\phi_0(r, \theta) = J_m(kr) \cdot e^{im\theta}. \tag{41}$$

We first choose $m_0 > k$ large enough to attain the skin effect (see Section 2.7.2), so that the Born approximation is valid to linearize the relationship between q and ψ. The resulting linear system is solved to obtain the visible, or observable, part of q, denoted here by q_{m_0}, around the rim of the disk scatterer.

Remark 3.7 *The solution of the linear system is of least-squares type, and indeed, is not in general a very good approximation to q_{m_0}. But the error made there will be dealt with later on. Therefore, for simplicity, we still denote the least-squares solution by q_{m_0}.*

Now, suppose we have found a forward model q_n which produces the prescribed measurements for incoming fields (41) with $m = m_0,\ m_0 - 1,\ \ldots, n$. Suppose further that the incoming fields penetrate to a depth at $r = r_n$ with $0 < r_n < 1$, and thus cover the annulus $A(r_n, 1 - r_n)$ (see (23)) bounded by the two circles $r = r_n$ and $r = 1$. We remark that q_n is the observable part of q in $A(r_n, 1 - r_n)$ which in general is not q restricted in $A(r_n, 1 - r_n)$. Now, we wish to obtain a forward model q_{n-1} that produces the prescribed measurements for incoming fields not only with $m = m_0,\ m_0 - 1,\ \ldots, n$, but also with $m = n - 1$.

Remark 3.8 *In general, q_n and q_{n-1} are not the same in $A(r_n, 1-r_n)$, but both look the same to the scattering experiment with incident waves (41) for $m = m_0,\ m_0 - 1,\ \ldots, n$. The "domain of influence", $A(r_n, 1-r_n)$, is actually not exactly an annulus, due to, for example, the bending of the ray tubes. Finally, there is a transition zone associated with each domain of influence, for near the inner boundary of $A(r_n, 1-r_n)$, the influence is weak, the scattering is also weak, and the relationship between q and ψ is essentially linear within the zone, as well as without the zone, in the area nearer the center of the disk scatterer.*

The recursive linearization procedure is as follows. The forward model q_n is used to solve the forward scattering problem (5) corresponding to the incoming fields (41) with $m = m_0,\ m_0-1,\ \ldots, n,\ n-1$. The resulting scattered fields $\check{\psi}_m$ satisfy the $m_0 - n + 2$ equations

$$\check{\psi}_m(x) = -k^2 \int_\Omega G_k(x,\xi) \cdot q_n(\xi) \cdot (J_m(k\rho) \cdot e^{im\alpha} + \check{\psi}_m(\xi)) \cdot d\xi, \tag{42}$$

with the vector $\xi = \rho \cdot (\cos\alpha, \sin\alpha)$.

Remark 3.9 *According to Remark 3.7, the forward model q_n may have non-negligible error in it; consequently, $\check{\psi}_m(x)$, for $m \geq n$ and at the measurement points x outside the scatterer, are not the same as the prescribed scattering data.*

By definition, q_{n-1} is the forward scattering model observable to the scattering experiment using incoming fields (41) with $m = m_0,\ m_0 - 1,\ \ldots, n,\ n-1$, and therefore

$$\psi_m(x) = -k^2 \int_\Omega G_k(x,\xi) \cdot q_{n-1}(\xi) \cdot [J_m(k\rho) \cdot e^{im\alpha} + \psi_m(\xi)] \cdot d\xi, \tag{43}$$

with $\psi_m(x)$, for $m \geq n-1$ and x outside the scatterer, the prescribed scattering data. Now, subtracting (42) from (43) for all $m = m_0,\ m_0 - 1,\ \ldots, n-1$, we obtain a system of essentially linear relationships between the perturbation

$$\delta q = q_{n-1} - q_n \tag{44}$$

and the perturbations

$$\delta\psi_m = \psi_m - \check{\psi}_m, \tag{45}$$

governed by $m_0 - n + 2$ Lippmann-Schwinger equations (with second order terms dropped):

$$\delta\psi_m(x) = -k^2 \int_\Omega G_k(x,\xi) \cdot [(J_m(k\rho) \cdot e^{im\alpha} + \check{\psi}_m(\xi))\delta q(\xi) + q_n \delta\psi_m(\xi)] d\xi. \tag{46}$$

Given q_n, $\check{\psi}_m$, for $m = m_0,\ m_0 - 1,\ \ldots, n-1$, which are indeed known, the integral equation can be solved to obtain a linear expression of $\delta\psi$, at the measurement points outside the disk scatterer, via δq. Denoting this linear procedure by $\mathcal{L}_m$, we have

$$\delta\psi_m = \mathcal{L}_m(\delta q). \tag{47}$$

The linear system can be solved to yield δq, and q_{n-1} can be obtained via (44).

The remarks made at the end of the preceding subsection apply to this layer-stripping scheme. Not all $m_0 - n + 2$ equations (46) need to be solved, namely, those with higher m may be dropped from the simultaneous linear system (47). The numerical implementations and results will be presented on a later date.

4 Generalizations

In this last section, we first make several general remarks on the extensions of the recursive linearization procedures to three dimensions, to case of limited aperture, and to similar inverse problems. We then extend the results to other inverse problems and more general ill-posed nonlinear problems.

1. The extension of the inversion methods to three dimensions is straightforward, but not practical if full-aperture data are used. The recursive linearization procedures will, in this case, cost $O(N^8)$ operations to reconstruct an $N \times N \times N$ wavelengths problem in three dimensions, which is utterly prohibitive. Although careful implementations of the methods will lead to an $O(N^6)$ procedure, no numerical experiments have been done.

2. This brings us to the question of reconstruction from a limited aperture. That turns out to be completely simple in the lights of discussion in Section 2.6, and of Observation 2.13 made there. For a limited aperture, one can, of course, only recover the observable part of the scatterer (see (25)). The nonlinear problem of inverse scattering with limited aperture is formulated and approached in exactly the same way as for the case of full aperture. Then, due to Observation 2.13, the perturbation analyses of Section 3 are still valid, and the nonlinear system breaks into a collection of linear systems where the non-observable part of the scatterer, or the corresponding perturbation which is not observable to the limited aperture, falls into the null space of the linear systems, and therefore can be dealt with the same way as the non-observable part of the scatterer to a full aperture is. The whole issue of inversion from a limited aperture is, in this sense, not a jot or tittle different from that of the full aperture. In many interesting applications where the sources and receivers are essentially co-located, the use of Gaussian beams (see Section 2.7.3) to recursively linearize the nonlinear problem seems attractive.

3. The application of the recursive linearization procedures can, obviously, be made to other types of inverse scattering problems associated with wave phenomena, whether acoustic, electromagnetic, or else. The forward model can be more realistic, and therefore, more complicated, than that governed by the simple Helmholtz equation (1), so far as the fundamental principles used in this paper are still valid in these environments.

4.1 Layer Stripping for Electrical Impedance Imaging

In this section, we discuss very briefly the advantages and applications of a single frequency inversion method to solving an inverse problem not associated with wave phenomena. Here we will take the problem of electrical impedance imaging as an example to examine the method presented in Section 3.2.

For inverse problems not related to wave phenomena, the frequency k, as a parameter on which perturbation analyses can be performed, is absent. Usually, the problem can be classified as of zero frequency or extremely low frequency, one frequency, imaginary frequency, or simply, no frequency. Electrical impedance imaging and induction tool prospecting are two examples of such a character, where the equations are strongly elliptical. Therefore, there are the so-called skin effects, both in the classical sense and in that defined in Section 2.7.2.

In the case of electrical impedance imaging, the incident (electrical potential) fields are of the form

$$\phi_0(r,\theta) = r^m \cdot e^{im\theta}, \tag{48}$$

inside a disk. As a result, the incoming field undergoes rapid decay for m in the order of 10. It is not hard to observe that the two functions r^{10} and r^{11} look similar, and that to a reasonable precision, both behave like a (half of a) delta function in any interval $[0, A]$. This means that the highest m we choose to start the recursive linearization (see Section 2.7.2) is on the order of 10. It also means that the number of parameters to be recovered, independent of the size of the scatterer, independent of the method used, is on the order of 10×10. In many applications, even such a limited resolution is difficult to achieve, and if attainable, would be extremely useful.

4.2 Nonlinear Ill-posed Problems

Inverse problems, such as those discussed above, generally belong to a somewhat wider class of problems – the nonlinear ill-posed problems. First, ill-posed linear problems we know very well how to solve, and are, in this sense, not interesting. Ill-posed nonlinear problems we know little about, and are often very interesting.

There is, however, one direction in which we may see more clearly. And that is, to a class of nonlinear ill-posed problems such as several types of inverse problems, ill-posedness is often not a curse, but a blessing. It helps linearize the nonlinear problem, and once linearized, the collection of linear problems can be treated the best to the limitations of the underlying physics to in turn deal with the ill-posedness – these are ill-posed linear problems after all.

Let us consider a general nonlinear ill-posed problem of the form

$$\mathcal{F}(x) = y, \tag{49}$$

with x, y "vectors" living in their functional spaces. By definition of ill-posedness, we can choose "a system of coordinates," or a proper representation, in which a perturbation to some components of x will cause no change in the observation y. These components are what we called the non-observable part of x, which it makes no sense to recover. Then there are components in x observable to variable degree. We say these modes live in the transition zone in that system of coordinates.

Often, the nonlinear system is a reformulation of some differential equation (or integral equation) where y is a collection of the solutions of the differential equation, and x represents other parameters in the equation, such as coefficients, initial values, or boundary conditions. Sometimes, the differential equation describes a simple mechanical process such as wave propagation, where the resulting nonlinearity in (49) has a simple structure. For example, we could say that the nonlinear relationship between q and ψ in (5) is quadratic, in the sense that the nonlinearity is a result of the product of the two functions. Our experience indicates that frequently where the nonlinear system is the most ill-posed, there the problem is most linear.

The recursive methods discussed in this paper can be viewed as procedures solving "initial value problems" associated with the nonlinear mapping $\mathcal{F}$. It is even possible to derive a system of ODE in k for the method presented in Section 2.4. In contrast, the nonlinear optimization and related methods, and the Newton iterations and related methods deal with and apply the entire nonlinear operator, or its derivatives, to successfully generated vectors.

References

[1] J. Sylvester and G. Uhlmann (1986), *A uniqueness theorem for an inverse boundary value problem in electrical prospection, Comm. Pure Appl. Math.* **17**, pp. 91–112.

[2] J. Somersalo, M. Cheney, D. Isaacson and E. L. Isaacson (1991), *Layer stripping: A Direct numerical method for impedance imaging,* Inverse Problems 7 899–926.

[3] J. Sylvester (1992), *A Convergent Layer Stripping Algorithm for the Radially Symmetric Impedance Tomography Problem,* Communications in *PDE*$\mathbf{17}$,No.12, pp.1955–1994.

[4] D. Colton and R. Kress (1983), *Inverse Acoustic and Electromagnetic Scattering Theory,* Wiley-Interscience Publication, New York.

[5] Y. Chen and V. Rokhlin (1995), *On the Riccati Equations for the Scattering Matrices in two dimensions,* Technical Report 1081, Department of Computer Science, Yale University.