Inverse Problems in Medical Imaging

Charles L. Epstein
Department of Mathematics
Laboratory for Structural NMR
University of Pennsylvania

What is an inverse problem?

- A basic problem in medical imaging is to reconstruct an image of something inside the human body from minimally invasive, non-destructive measurements
- The measurements are related to the quantities of interest by a mathematical model, which usually describes how the "unknown" system would produce the measured values
- The basic "inverse problem" is to determine the system from sufficiently many measurements

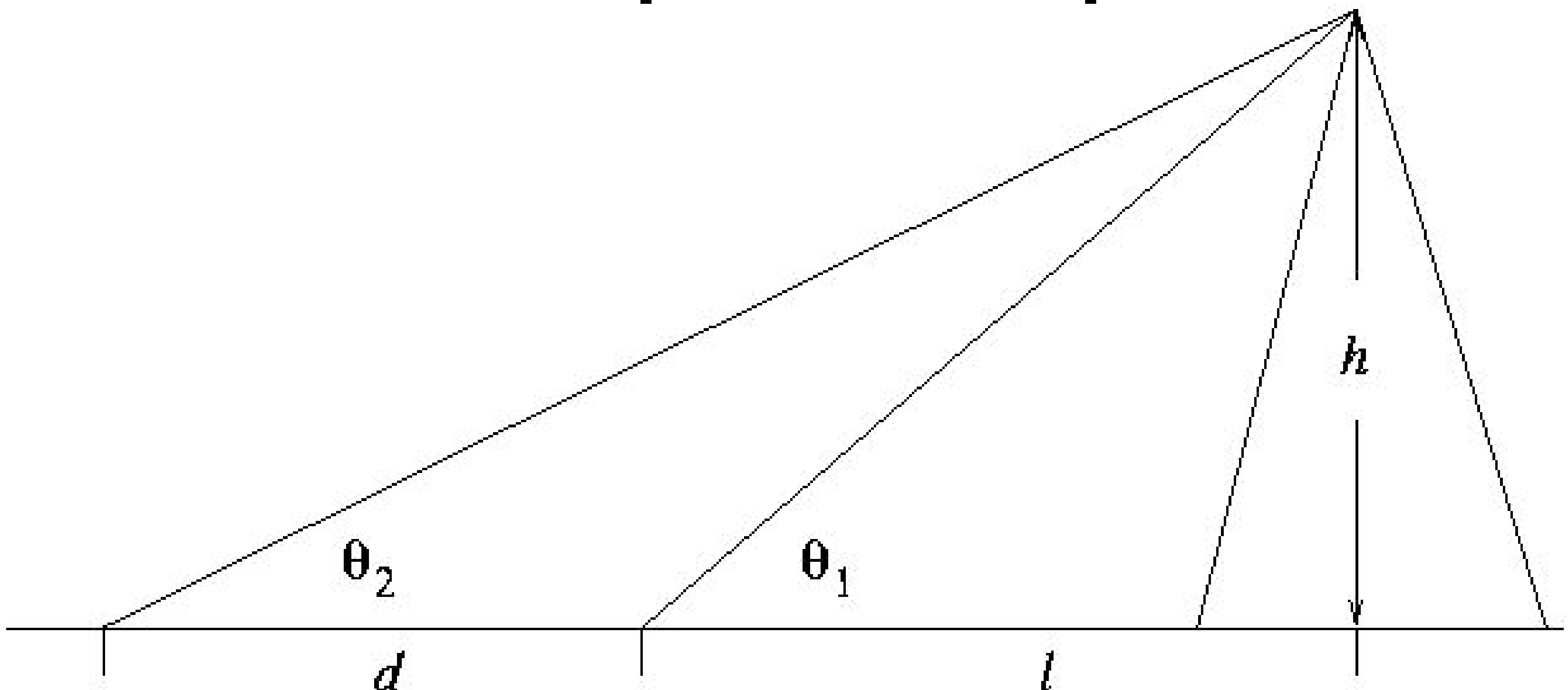
A mathematical description

- The system is specified by state variables X.
- The measurements Y are functions of the state variables, Y=A(X) (otherwise the state variables don't specify the state of the system).
- In many cases the map A: X ---> Y is linear, so the inverse problem starts out as the problem of inverting A.

Steps in the analysis of an inverse problem (the ideal)

- Uniqueness: Decide which measurments Y suffice, in principle, to determine X
- Reconstruction: From an exact inversion algorithm B to find X perfect data Y. This sometimes involves characterizing the range of A, that the set of possible measurements.
- Practical implementation: Find a stable, accurate approximation to B that can be applied to a finite, noisy set of measurements.

A simple example



- We can measure θ_1, θ_2 and would like to determine *h* and *l*.
- We use the relation $\tan\theta = \frac{h}{l}$ and find that

$$h = \frac{d\tan\theta_1\tan\theta_2}{\tan\theta_1 - \tan\theta_2} \qquad l = \frac{\tan\theta_1 - \tan\theta_2}{d\tan\theta_2}$$

Models are useful for estimating sensitivity to errors

- Angles close to 90 degrees lead to instability in the predictions.

$$\frac{\delta h}{h} = \delta\theta_1 \left(\frac{2}{\tan\theta_2 \sin 2\theta_1} - 1 \right) + \delta\theta_2 \left(\frac{2}{\tan\theta_1 \sin 2\theta_2} + 1 \right) + O(\delta\theta_1^2 + \delta\theta_2^2)$$

The Radon transform and filtered backprojection

$$R\rho(\omega, s) = \int_{\langle (x,y),\omega\rangle = s} \rho dl$$

$$\rho(x, y) = c_2 \int_{S^1} -is\partial_s \mathcal{H}_s R\rho(\omega, \langle (x, y), \omega\rangle) d\omega$$

$$\rho(x_m, y_l) \approx \frac{d}{2(M+1)} \sum_{k=0}^{M} \sum_{j=-N}^{N} R\rho(\langle (x_m, y_l), \omega(k\Delta\theta)\rangle, jd)\phi(\langle (x_m, y_l), \omega(k\Delta\theta)\rangle - jd)$$

The filtered backprojection formula is very nice because it makes sense for very general data and can be rationally approximated.

General mathematical structure of inverse problems

- There are two general types
- A(X)=Y or A(X,B,Y)=0
- In the first type X is the state and Y are measurements and it is just a matter of inverting a map.
- In the second type X are known inputs, Y are known outputs and B parametrizes the system. The problem is to determine B from sufficiently many input, output pairs

Linear examples I

- We consider the first type of problem, with A and linear map, so we want to solve AX=Y for X. In the end, we're always working with a finite dimensional problem...for simplicity let's assume A is invertible, so

$$\boldsymbol{X} = \boldsymbol{A}^{-1}\boldsymbol{Y}$$

- But....while A may be invertible it is frequently ill-conditioned.

Conditioning

- The condition number of A is defined to be

$$C_A = \frac{\max_{x\neq 0} \|Ax\|}{\min_{x\neq 0} \|Ax\|} = \|A^{-1}\| \|A\|$$

- Since A is usually a finite dimensional approximation to an operator in a Hilbert space the condition number often grows unboundedly as the dimension of the space grows.

- If δY is the uncertainty in the measurements, then δX, the uncertainty in X satisfies

$$\frac{\|\delta X\|}{\|X\|} \asymp C_A \delta Y$$

Notation

- We use the notation $a \asymp b$ to mean that a could be as large as b.

Noise versus resolution, the SVD

- Because measurements are noisy, we are forced to limit the resolution to obtain stable algorithms. This is easy to see in terms of the singular value decomposition (SVD).

$$\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^t$$

- The matrices U,V are unitary and

$$\boldsymbol{\Sigma} = \mathrm{diag}(\boldsymbol{\sigma}_1, \ldots, \boldsymbol{\sigma}_N)$$

- The $\sigma_1 \geq \cdots \geq \sigma_N$ are the singular values, and the condition number satisfies $$\boldsymbol{C_A} = \frac{\boldsymbol{\sigma}_1}{\boldsymbol{\sigma}_N}$$

Resolution vs. Noise

- The SVD allows us to express X in terms of the data:

$$U = (u_1, \ldots, u_N) \quad V = (v_1, \ldots, v_N) \qquad X = \sum_{j=1}^{N} \frac{\langle u_j, Y \rangle}{\sigma_j}$$

- The smaller singular values usually correspond to singular vectors with more oscillation, representing higher resolution in the solution:

$$v_j \approx [1, e^{\frac{2\pi i j}{N}, \ldots, e^{\frac{2\pi i j (N-1)}{N}}}]$$

- Because of noise we need eliminate "small", $\sigma_j \leq \sigma_{\min}$
- singular values....this limits the resolution in X.

$$X \approx \sum_{\{\sigma_j > \sigma_{\min}\}} \frac{\langle u_j, Y \rangle}{\sigma_j}$$

Linear examples, II

(A+B)X=Y

- Now we consider determining B from sufficiently many input-output pairs (X_j, Y_j)
- In fact, we can suppose that the inputs are arranged in a matrix, X, which is unitary so that B is given by $B = YX^{-1} - A$
- Usually B is a small perturbation of A, which can be taken to mean $\|B\| < \sigma_N(A)$
- This implies that if δX is the uncertainty in the inputs, then the uncertainty in the outputs is $\delta Y \approx A\delta X$ and so

$$\frac{\|\delta B\|}{\|B\|} \approx C_A \delta X$$

Conditioning, the good, the bad and the ugly

- Let H be a Hilbert space and A:H $\Longrightarrow$ H, a bounded operator.
- (Good) A is well conditioned if A^{-1} is bounded
- (Bad) A is mildly ill-conditioned if $\sigma_j(A) \geq \frac{M}{j^m}$
- (Ugly) A is severely ill-conditioned if

$$\sigma_j(A) = O(j^{-m}) \text{ for all } m$$

Modalities and their inverse problems

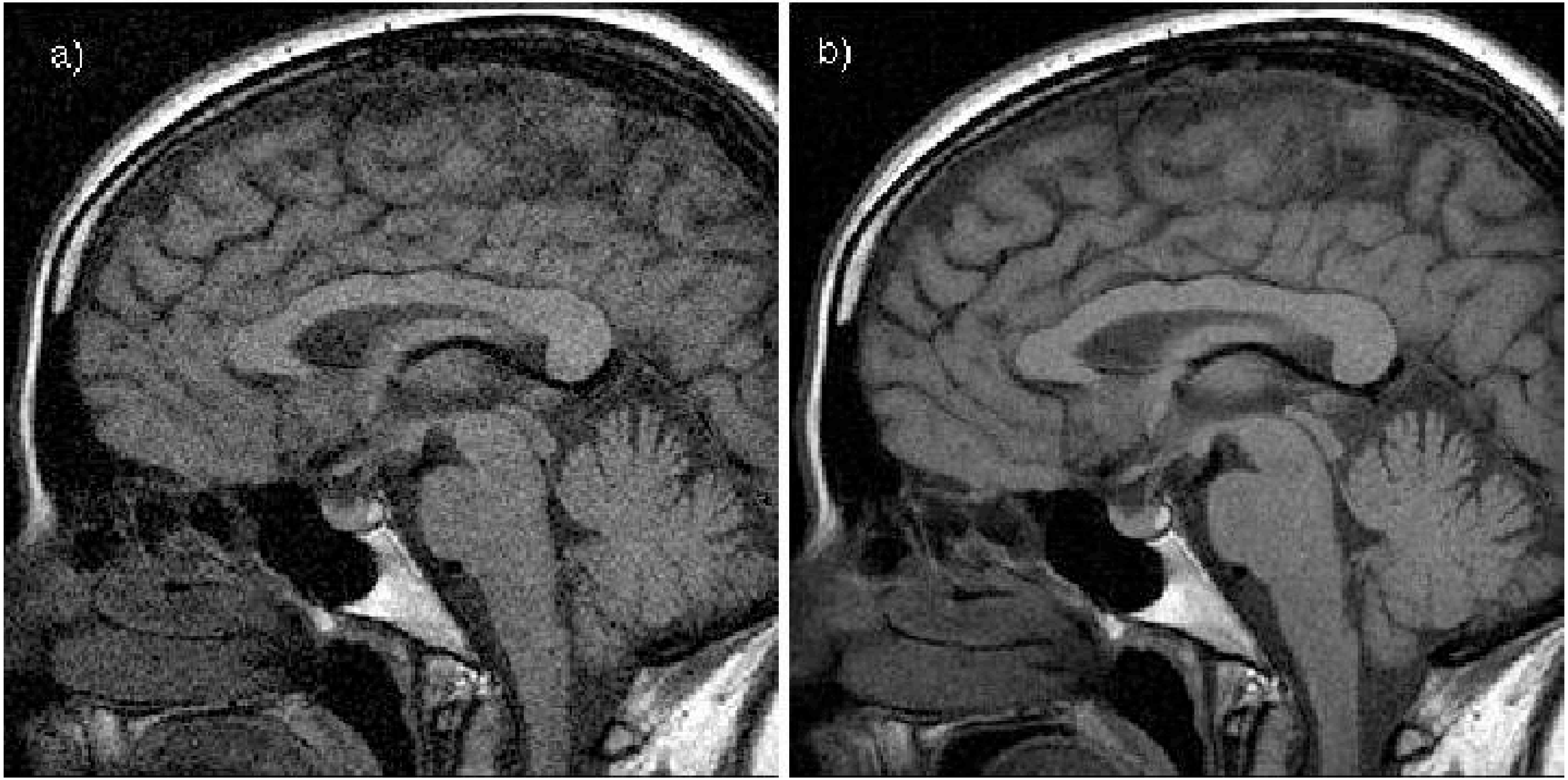
- In many ways the best case scenario is represented by MRI. The inverse problem is simply inversion of the Fourier transform, the measurements are modeled as

$$\hat{\rho}_j = \int_{-\delta}^{\delta} \psi(t - t_j) \left[\int_D \rho(x) e^{-2\pi i k(t)\cdot x} e^{-\frac{t+\tau}{T_2(x)}} dx + n_j(t) \right] dt$$

- Here ψ models the receiver and n_j is a white noise process.

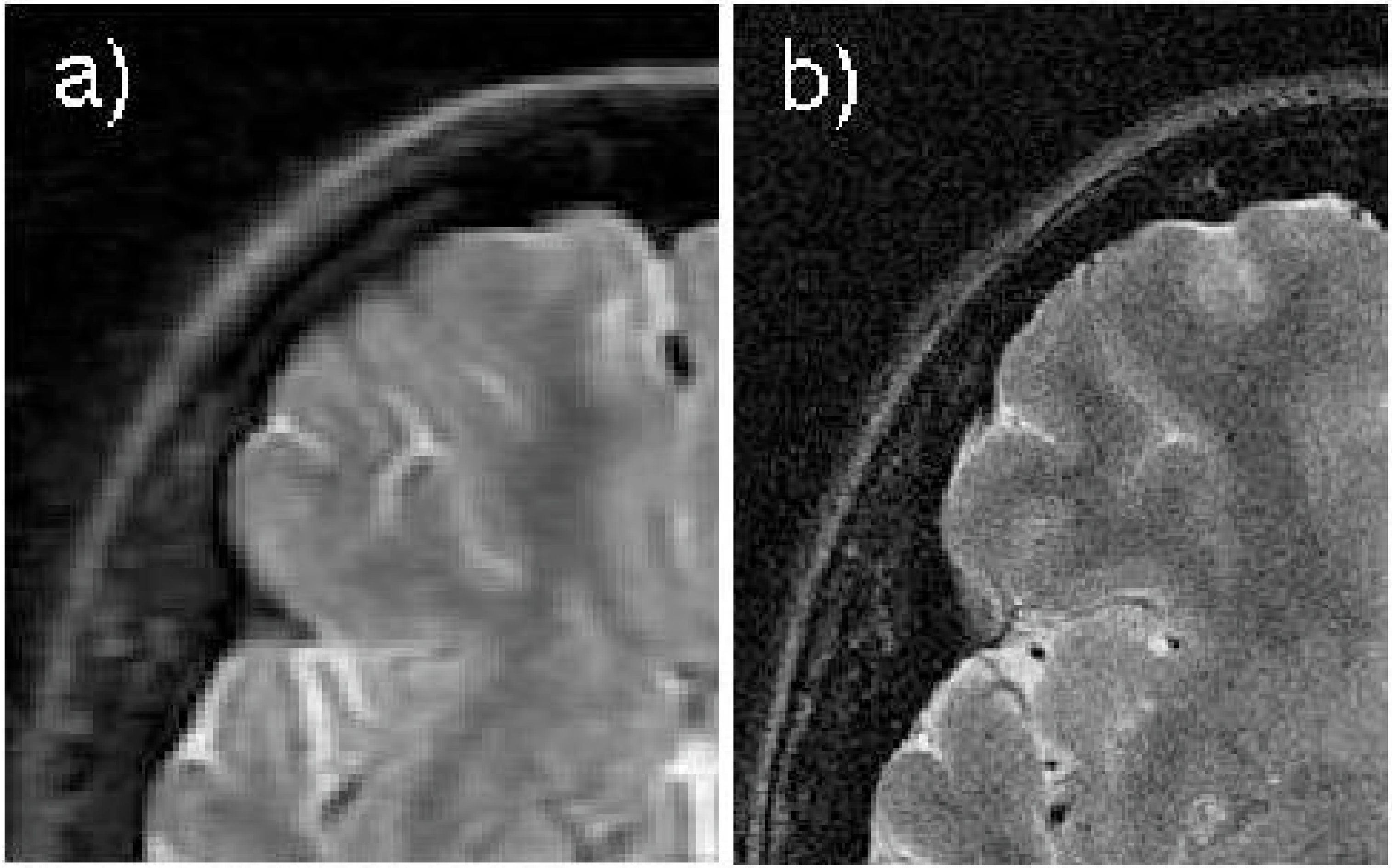
- Noise and the exponential decay impose effective limits on the resolution, even though the basic operator is unitary and hence well conditioned.

MR image 1



- MR image showing the effects of noise.

MR image 2



- MR images showing the effect of the maximum frequency sampled on resolution.

X-Ray CT

- The measurement is modeled as an averaged Radon transform

$$R\rho_{jk} = \int_{-\delta}^{\delta} \psi(s - s_j) \left[\int_{\{\langle (x,y), \omega_k \rangle = s\}} \rho dl \right] ds$$

- It can be interpreted as the Radon transform of a smoothed function.

$$R\rho_{jk} = R[\rho * \tilde{\psi}](s_j, \omega_k),$$

- The inverse is mildly ill-posed, the inverse involves taking a derivative of the measurements:

$$\rho(x, y) = C \int_{S^1} -i[\partial_s \mathcal{H}_s R\rho](\langle (x, y), \omega \rangle, \omega) d\omega$$

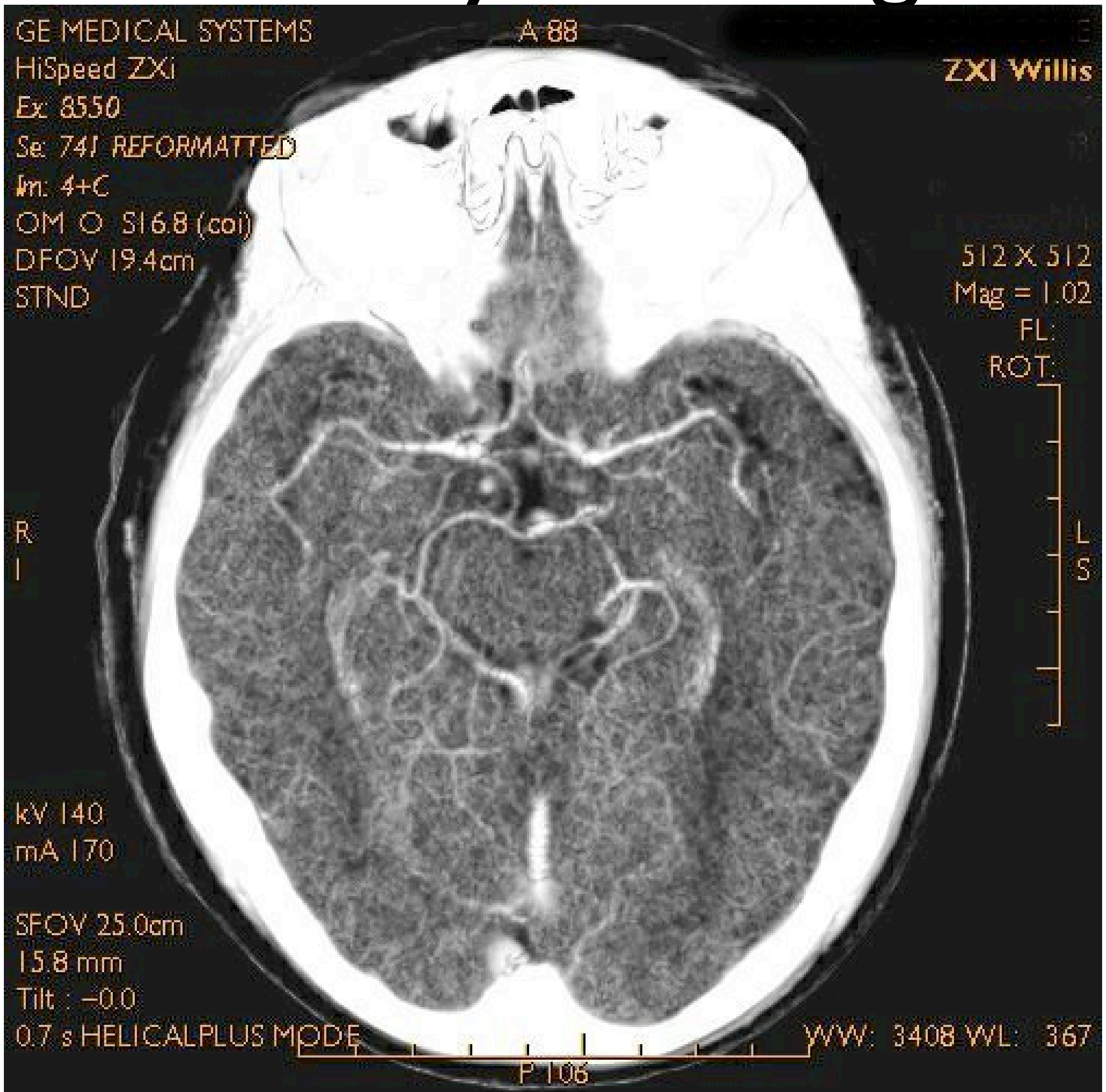
Reconstruction in X-Ray CT

- To obtain a stable reconstruction one needs to cut-off the data in Fourier space, this limits resolution.
- The SNR is proportional to the fourth power of the radiation dosage, so the resolution is limited by patient safety considsrations.

3d CT-imaging

- After many fallow years the introduction of cone beam machines, with many detectors, has lead to a significant renaissance of interest in 3d-reconstruction algorithms and problems in integral geometry connected to them.
- The problem of stable reconstruction with partial data sets remains largely unsolved and important....due to patient safety considerations.

A X-Ray CT image

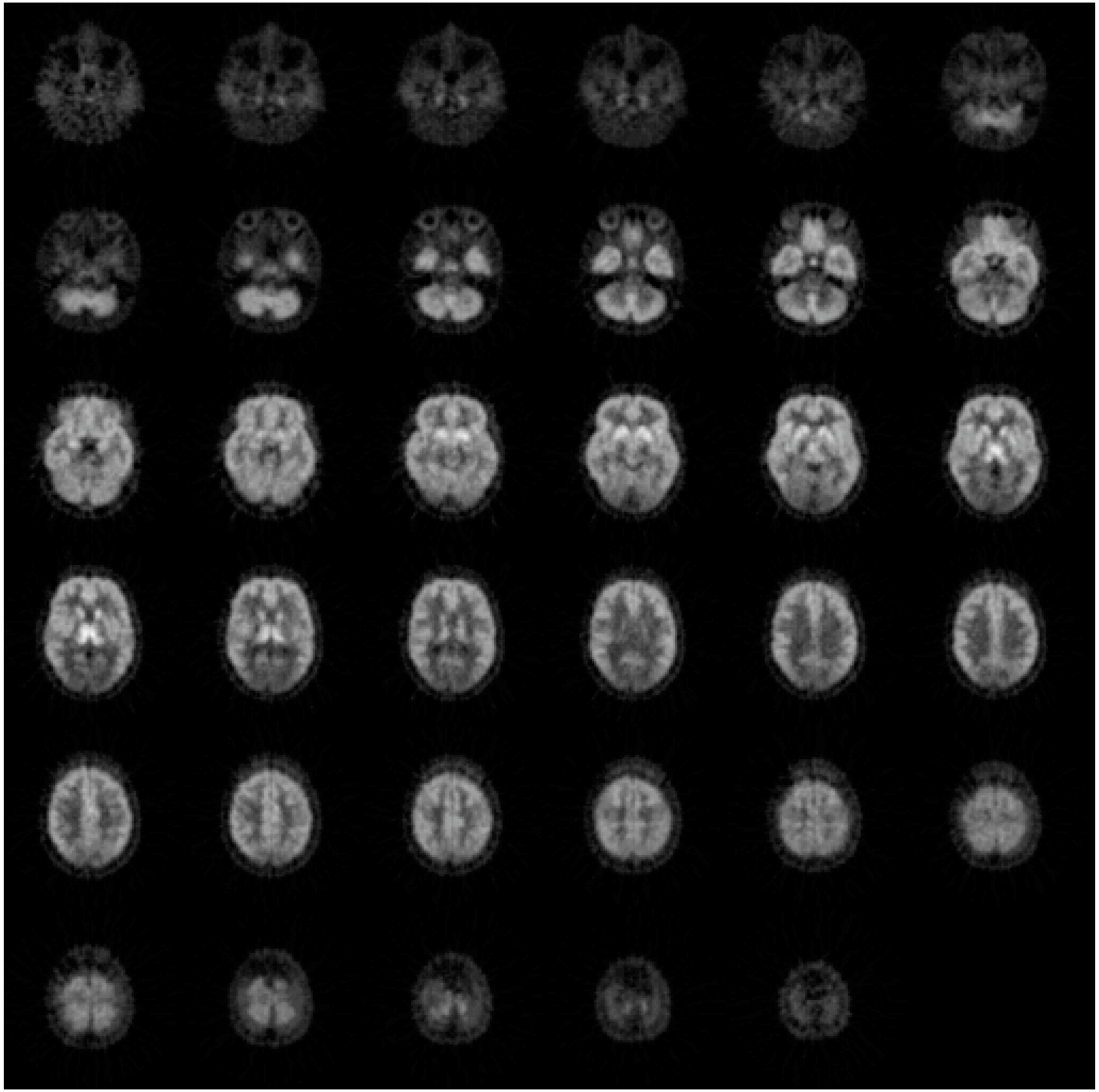


- High resolution, but poor soft tissue contrast.

Positron Emission Tomography

- In principle, the model for PET is the same as that for X-Ray CT, however, there is so much noise in the measurements, that this continuum model is not adequate and a less structured inversion method is usually employed.
- The images have much less resolution.
- Using a generic linear or nonlinear optimzation inverse algorithm is typical in problems, that for one reason or another lack a good continuum model.

PET image



- They look good, but are very small!

Ultrasound, EIT, DOT

- In principle, all these modalities are governed by similar models and are in essence inverse scattering problems.
- In ultrasound one can use a very crude model to obtain usable images, but there is very little mathematical processing and it is not possible to do much signal averaging to reduce measurement noise.
- Inverse scattering problems are severely ill-posed.

Inverse scattering

- A relatively simple example is provided by the Helmholtz equation. We illuminate the unknown object (modeled by q(x)) with a plane wave of frequency $2\pi k$ and wavelength:

$$\lambda = \frac{1}{k} \qquad u = u_s(\omega, k; x) + e^{2\pi i k \omega \cdot x}$$

- The physics is modeled by

$$(\Delta + q + (2\pi k)^2) u_s = -q e^{2\pi i k \omega \cdot x} \text{ where}$$

$$r|u_s(x)| < C \text{ and } r(\frac{u_s}{r} - iku_s) \to 0 \text{ as } r \to \infty.$$

Inverse scattering, II

- Measured data are the scattered waves which we encode as an operator on a Hilbert space $\Lambda_q : H_1 \to H_2$.
- The operator depends continuously on the data:

$$\|\Lambda_{q_1} - \Lambda_{q_2}\|_{H_1,H_2} \leq C\|q_1 - q_2\|_{L^\infty}$$

- The potential satisfies a very unfavorable estimate:

$$\|q_1 - q_2\|_{L^\infty} \leq \frac{M}{\left[\log\left(1 + \|\Lambda_{q_1} - \Lambda_{q_2}\|_{H_1,H_2}^{-1}\right)\right]^m}$$

The bad news

- The estimate on the previous slide says something like this: in order to determine q(x) with ONE decimal place of accuracy, we need to measure Λ_q with TEN decimal places of accuracy
- But this is not the end of the story.

The Rayleigh limit

- In the 17th-19th centuries a great deal of effort was expended to improve the resolution of microscopes and telescopes.
- In the late 19th century, Abbe and Rayleigh discovered that there is a limitation on the resolution, even if the optics are perfect. It follows from diffraction theory that the maximum resolution depends on the wavelength of the illumination:

$$\Delta x > c\lambda$$

- Methods exist to get beyond the Rayleigh...but not very far.

Evanescent waves

- A very similar effect is apparent from the plane wave expansion to a solution to the free space Helmholtz equation:

$$u(x,y,z) = \int\limits_{\xi_1^2+\xi_2^2<\lambda^2} \tilde{u}(\xi_1,\xi_2,0)e^{2\pi i(x\xi_1+y\xi_2)}e^{2\pi i z\lambda^{-1}\sqrt{1-\lambda^2(\xi_1^2+\xi_2^2)}}d\xi_1 d\xi_2 +$$

$$\int\limits_{\xi_1^2+\xi_2^2\geq\lambda^2} \tilde{u}(\xi_1,\xi_2,0)e^{2\pi i(x\xi_1+y\xi_2)}e^{-2\pi z\lambda^{-1}\sqrt{\lambda^2(\xi_1^2+\xi_2^2)-1}}d\xi_1 d\xi_2.$$

The second integral contains the high frequency information in u along z=0 and it decays exponentially for positive z. These are the *evanescent* waves, and this expression explains why it is so difficult to beat the Rayleigh limit.

The Born Approximation

- The measurement in a scattering situation is the scattering operator :
- $$s_q(\omega, k, \eta) = \lim_{r \to \infty} r e^{-2\pi i k r} u_s(\omega, k; r\eta)$$
- If the support of q is large compared to and q is small enough then we can use the Born approximation:

$$s_q(\omega, k, \eta) \approx \hat{q}(2\pi k(\omega - \eta))$$

- This shows that, if we stay within the Rayleigh limit, then in principle we can do fairly well.

Recent work

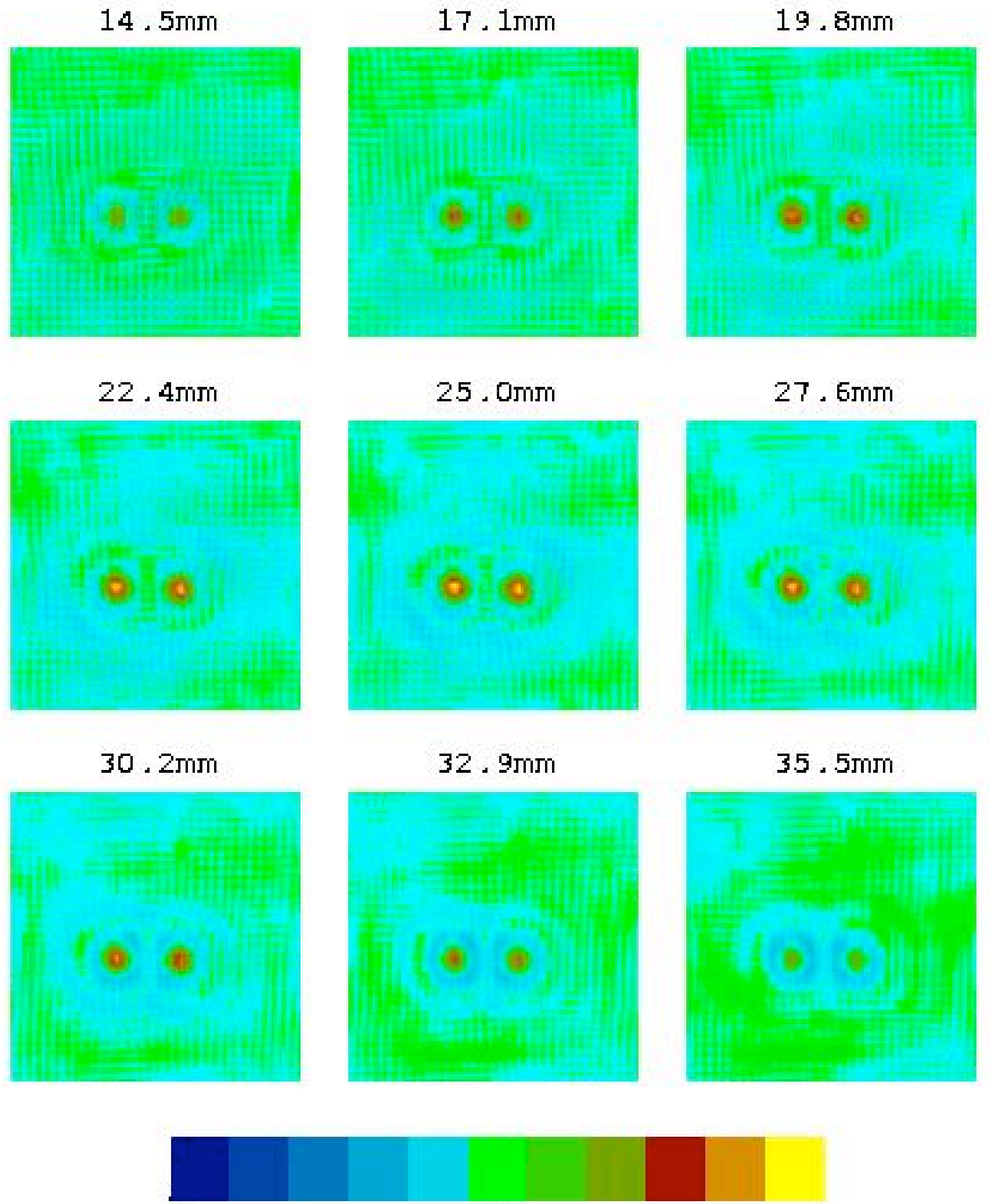
- In a recent paper Mike Taylor gave the beginning of a reconstruction method for the acoustical scattering problem that showed that, if one remained within the Rayleigh, then one should be able to obtain a stable algorithm. However, there are significant problems in the non-linear part of the process

- Roman Novikov has given a non-linear algorithm that gives a stable Rayleigh limited reconstruction for the potential in the Helmholtz equation.

Beating the Rayleigh limit

- In many problems in Diffuse Optical Tomography, all of the measurements consist of evanescent waves. By using a carefully controlled experimental design allowing for vast oversampling (~10^4 times) and usage of an explicit SVD to control the noise in the reconstruction, John Schotland et al. have obtained better than expected images using this very problematic modality.

DOT images



Prospects, I

- Once a physical measurement is decided upon then mathematics provides the tools to relate the measurements to the state of the system
- The model then gives a fairly precise idea of what is reasonably attainable, given the physical realities of the measurement process: feasible datasets, noise, relaxation and signal strength.
- Many interesting and important inverse problems are largely unsolved, but mathematicians should direct their efforts towards potentially useful modalities.
- It may be best to change the "rules of the game."

Prospects, II

- Many of the physical phenomena used in imaging modalities (especially MRI, ultrasound) are rich in new possibilities, Diffuison Tensor Imaging, Multiple quantum coherence, different types of waves in ultrasound....
- I see that the best chance for sugnificant progress lies in close collaboration among mathematicians, physicists, engineers and physicians.
- A big challenge is to interest mathematicians