Lecture 24

Exemplary Inverse Problems including Vibrational Problems

Syllabus

Lecture 01	Describing Inverse Problems
Lecture 02	Probability and Measurement Error, Part 1
Lecture 03	Probability and Measurement Error, Part 2
Lecture 04	The L_2 Norm and Simple Least Squares
Lecture 05	A Priori Information and Weighted Least Squared
Lecture 06	Resolution and Generalized Inverses
Lecture 07	Backus-Gilbert Inverse and the Trade Off of Resolution and Variance
Lecture 08	The Principle of Maximum Likelihood
Lecture 09	Inexact Theories
Lecture 10	Nonuniqueness and Localized Averages
Lecture 11	Vector Spaces and Singular Value Decomposition
Lecture 12	Equality and Inequality Constraints
Lecture 13	L_1 , L_∞ Norm Problems and Linear Programming
Lecture 14	Nonlinear Problems: Grid and Monte Carlo Searches
Lecture 15	Nonlinear Problems: Newton's Method
Lecture 16	Nonlinear Problems: Simulated Annealing and Bootstrap Confidence Intervals
Lecture 17	Factor Analysis
Lecture 18	Varimax Factors, Empircal Orthogonal Functions
Lecture 19	Backus-Gilbert Theory for Continuous Problems; Radon's Problem
Lecture 20	Linear Operators and Their Adjoints
Lecture 21	Fréchet Derivatives
Lecture 22	Exemplary Inverse Problems, incl. Filter Design
Lecture 23	Exemplary Inverse Problems, incl. Earthquake Location
Lecture 24	**Exemplary Inverse Problems, incl. Vibrational Problems**

Purpose of the Lecture

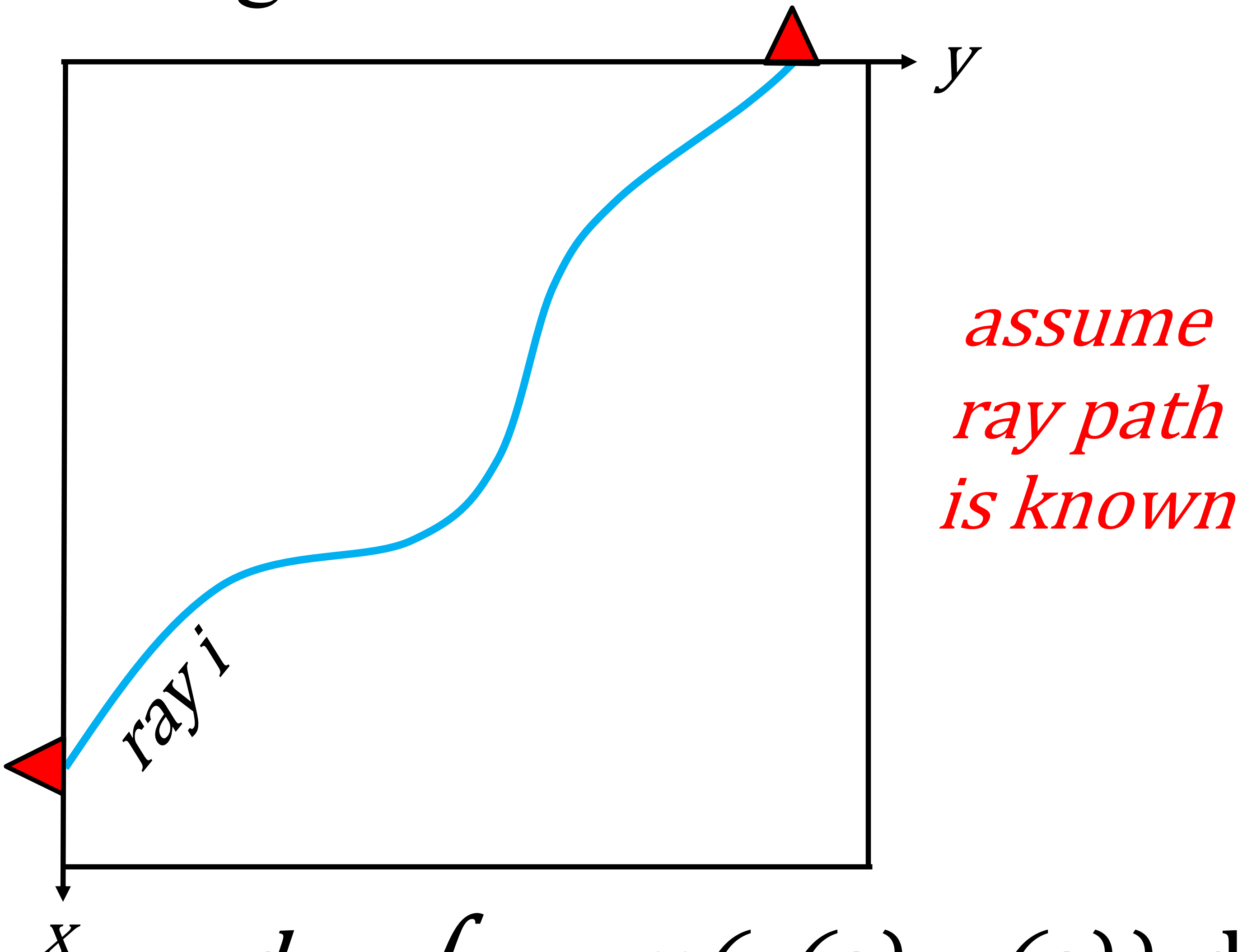
solve a few exemplary inverse problems

tomography
vibrational problems
determining mean directions

Part 1

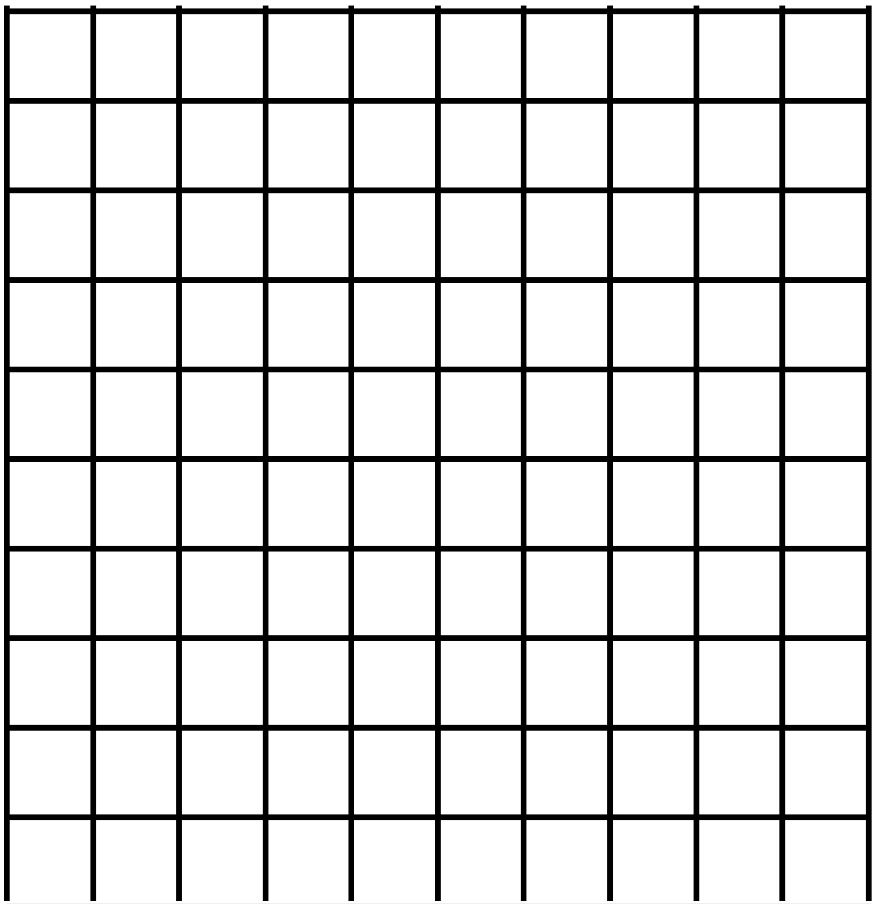
tomography

tomography:
data is line integral of model function



$$d_i = \int_{\text{ray } i} m(x(s), y(s))\, \mathrm{d}s$$

discretization:
model function divided up into M pixels m_j



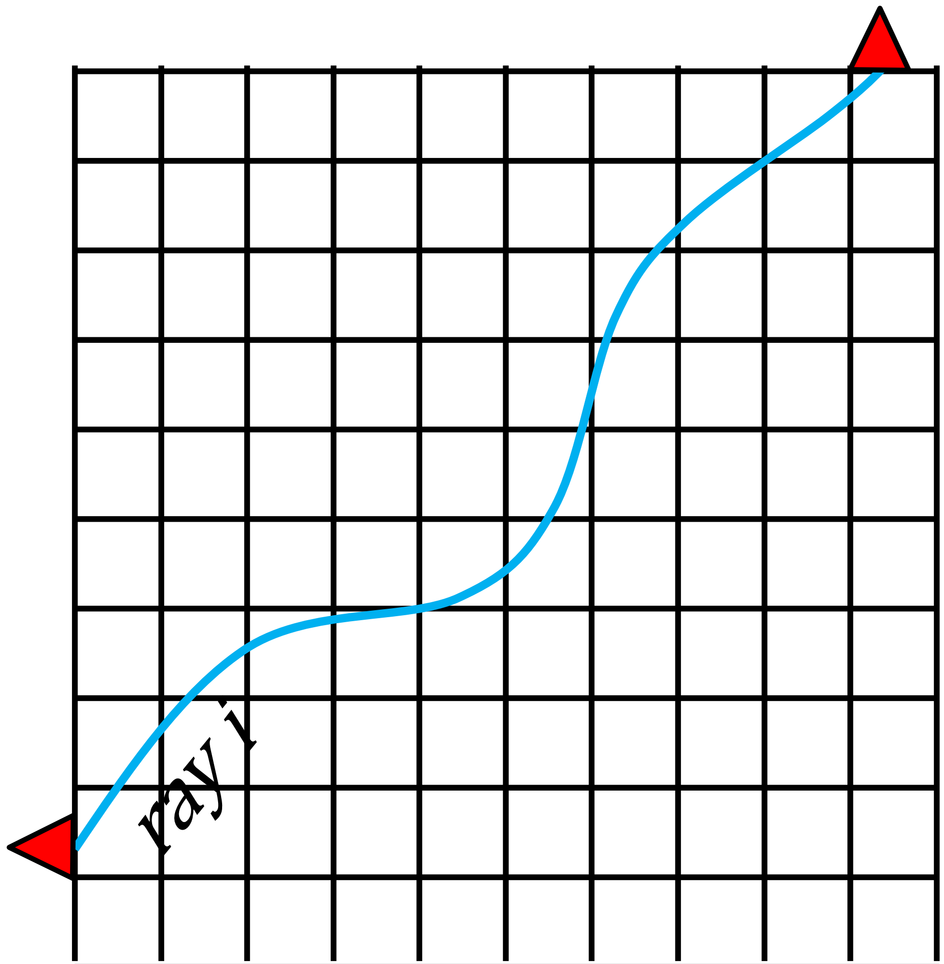
data kernel

G_{ij} = length of ray i in pixel j

data kernel

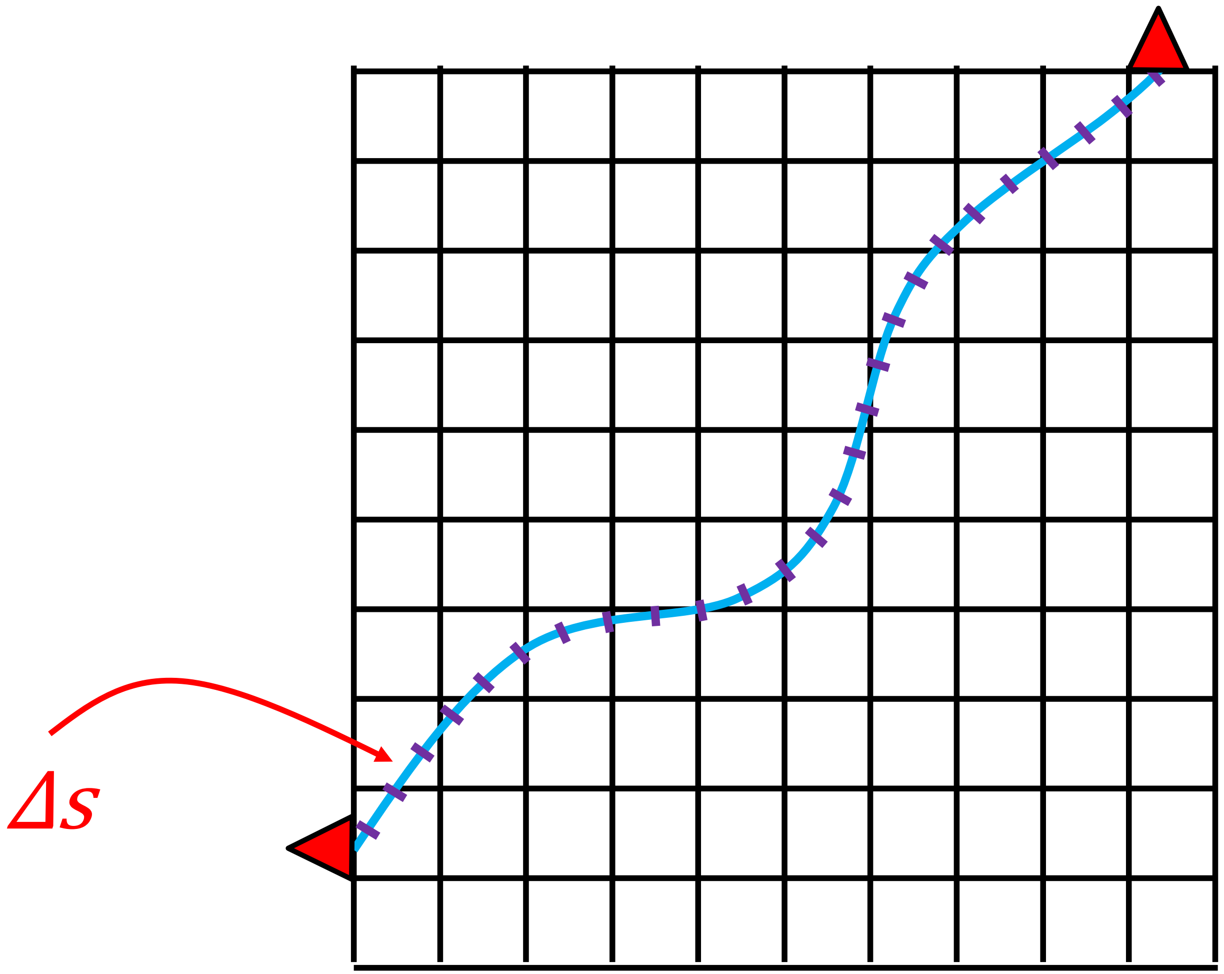
G_{ij} = length of ray i in pixel j

here's an easy, approximate way to calculate it start with **G** set to zero



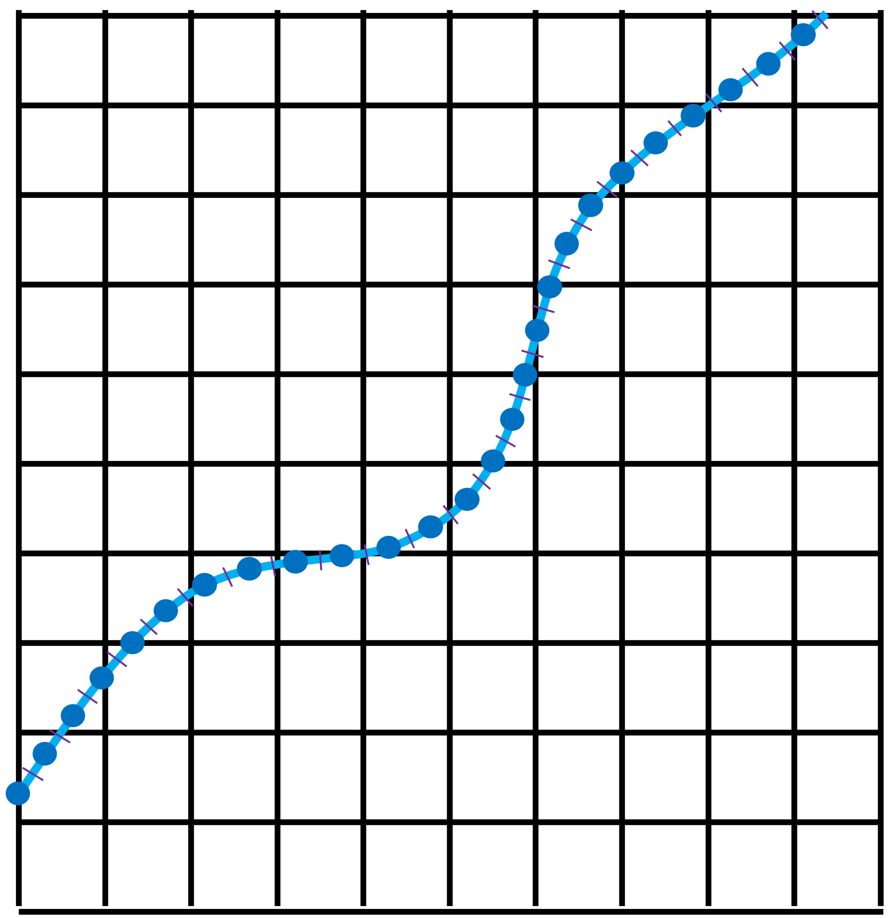
then consider each ray in sequence

divide each ray into segments of arc length Δs



and step from segment to segment

determine the pixel index, say *j*, that the *center* of each line segment falls within



add Δs to G_{ij}

repeat for every segment of every ray

You can make this approximation indefinitely accurate simply by decreasing the size of Δs

(albeit at the expense of increase the computation time)

Suppose that there are $M=L^2$ voxels

A ray passes through about L voxels

G has NL^2 elements
NL of which are non-zero

so the fraction of non-zero elements is
$1/L$

hence
G is very sparse

In a typical tomographic experiment

some pixels will be missed entirely

and some groups of pixels will be sampled by only one ray

In a typical tomographic experiment

some pixels will be missed entirely

the value of these pixels is completely undetermined

and some groups of pixels will be sampled by only one ray

only the average value of these pixels is determined

hence the problem is mixed-determined
(and usually $M>N$ as well)

so

you must introduce some sort of a priori information to achieve a solution

say

a priori information that the solution is small

or

a priori information that the solution is smooth

Solution Possibilities

1. Damped Least Squares (implements smallness):

   Matrix **G** is sparse and very large

   use **`bicg()`** with damped least squares function

2. Weighted Least Squares (implements smoothness):

   Matrix **F** consists of **G** plus

   second derivative smoothing

   use **`bicg()`** with weighted least squares function

Solution Possibilities

1. Damped Least Squares:

   Matrix **G** is sparse and very large

   use **`bicg()`** with damped least squares function

test case has very good ray coverage, so smoothing probably unnecessary

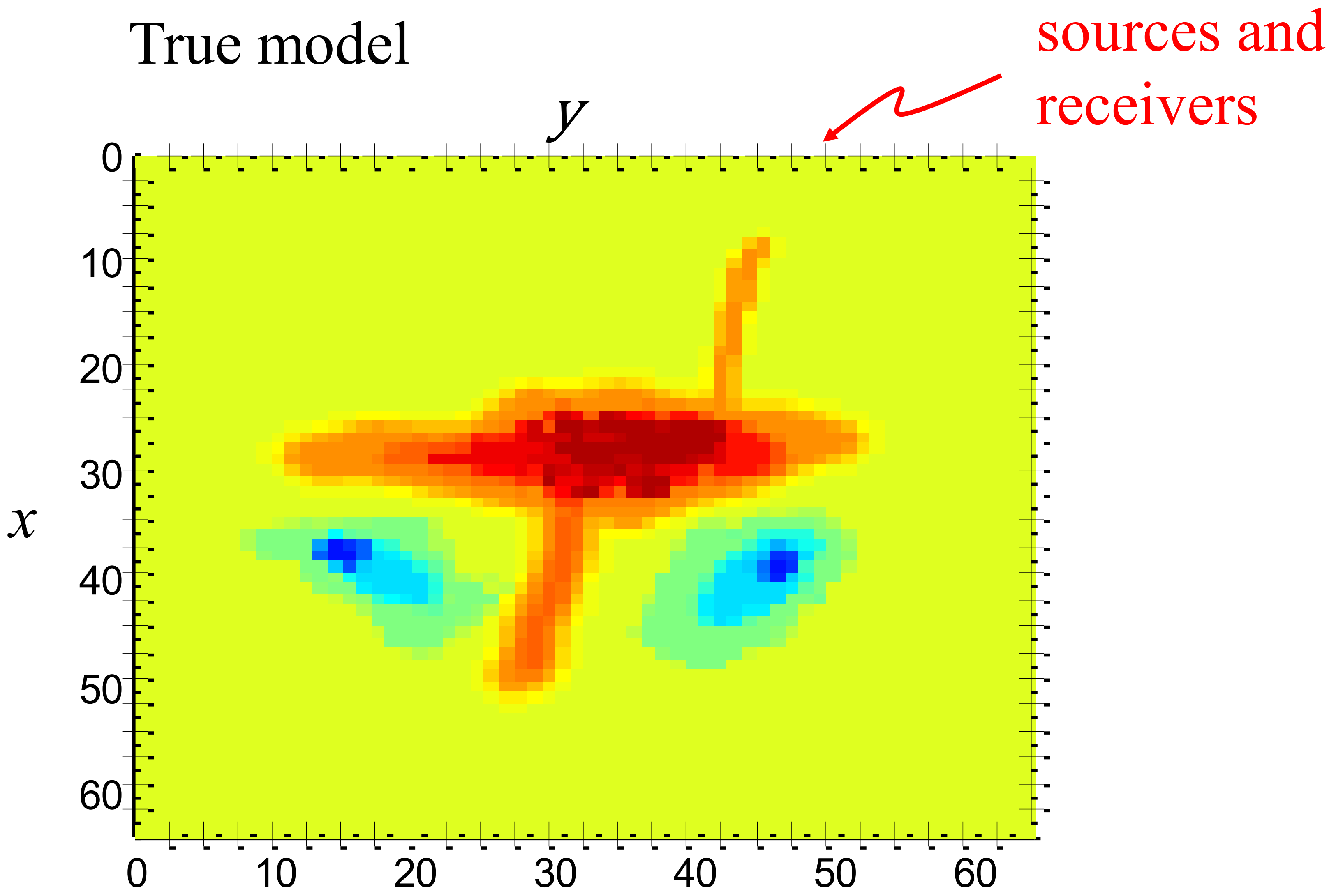
2. Weighted Least Squares:

   Matrix **F** consists of **G** plus

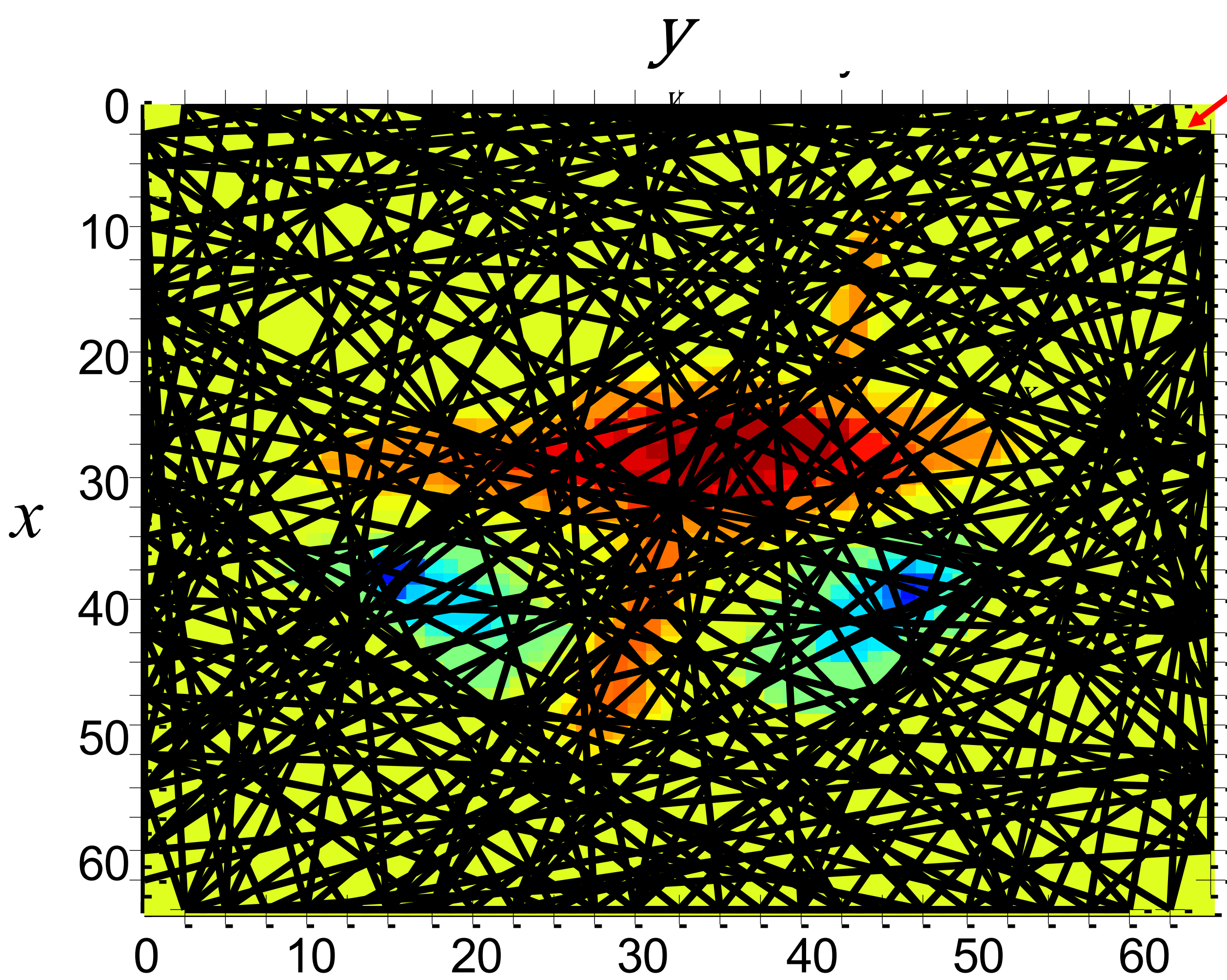
   second derivative smoothing

   use **`bicg()`** with weighted least squares function

True model



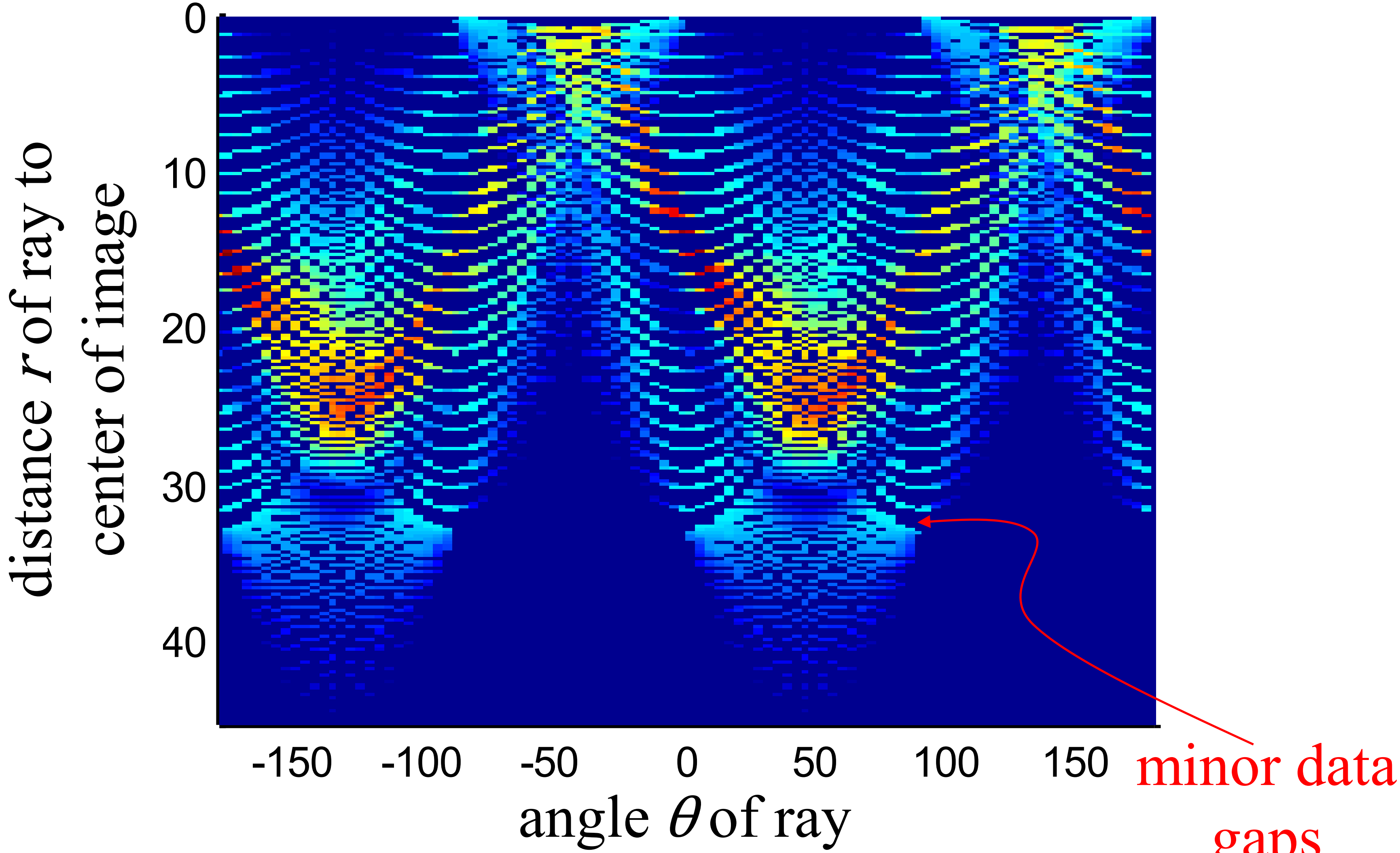
Ray Coverage



just a “few” rays shown

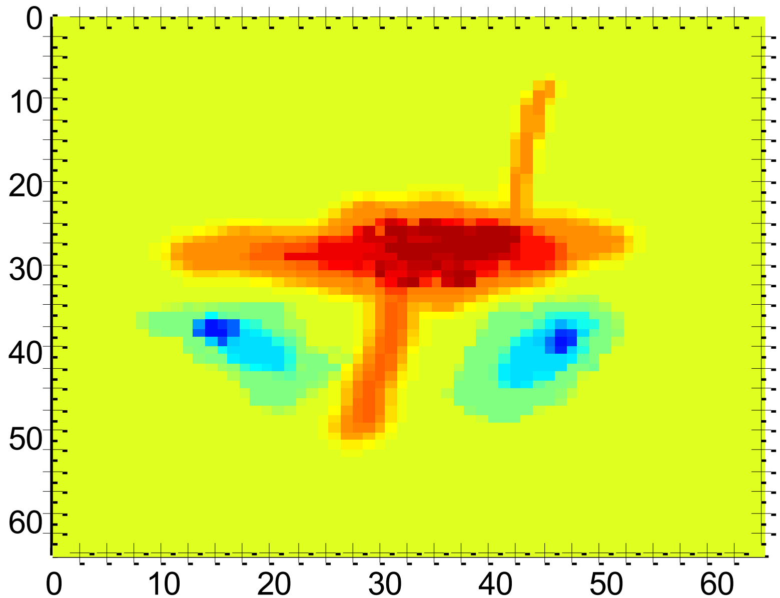
else image is black

Data, plotted in Radon-style coordinates

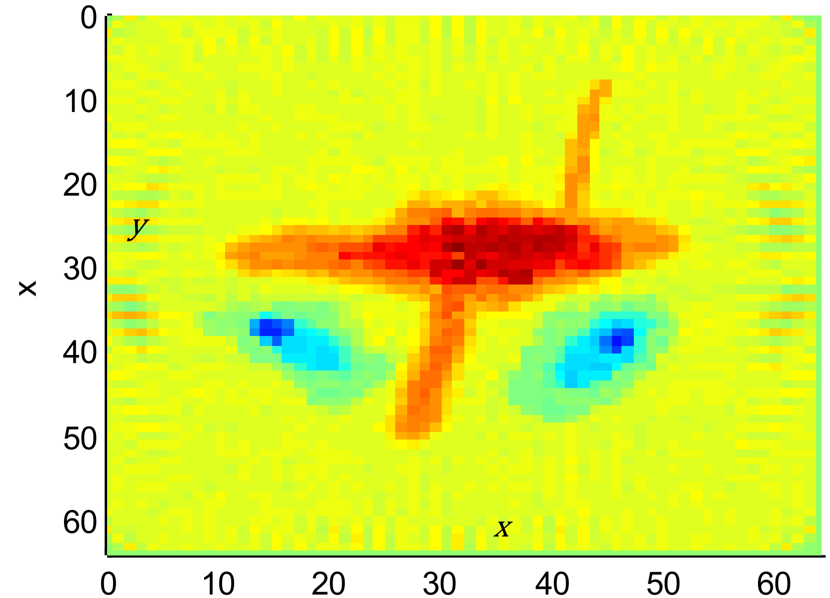


Lesson from Radon's Problem:
Full data coverage need to achieve exact solution

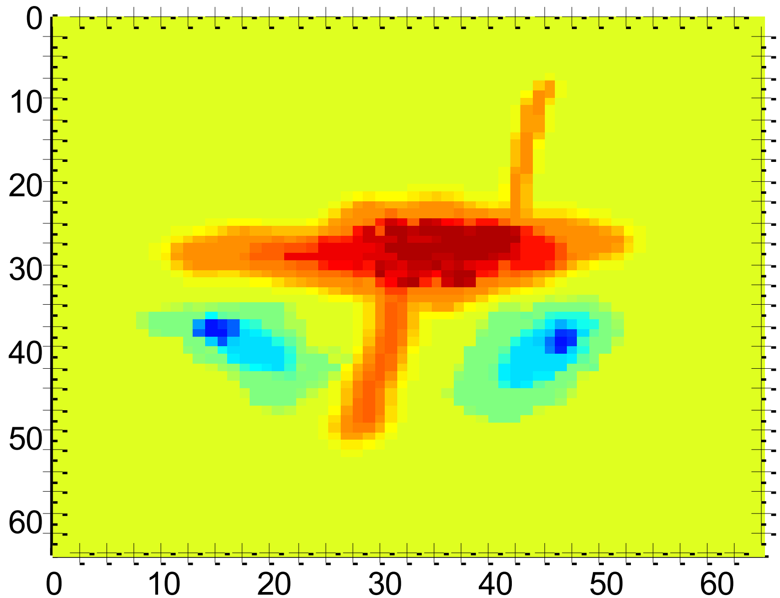
True model



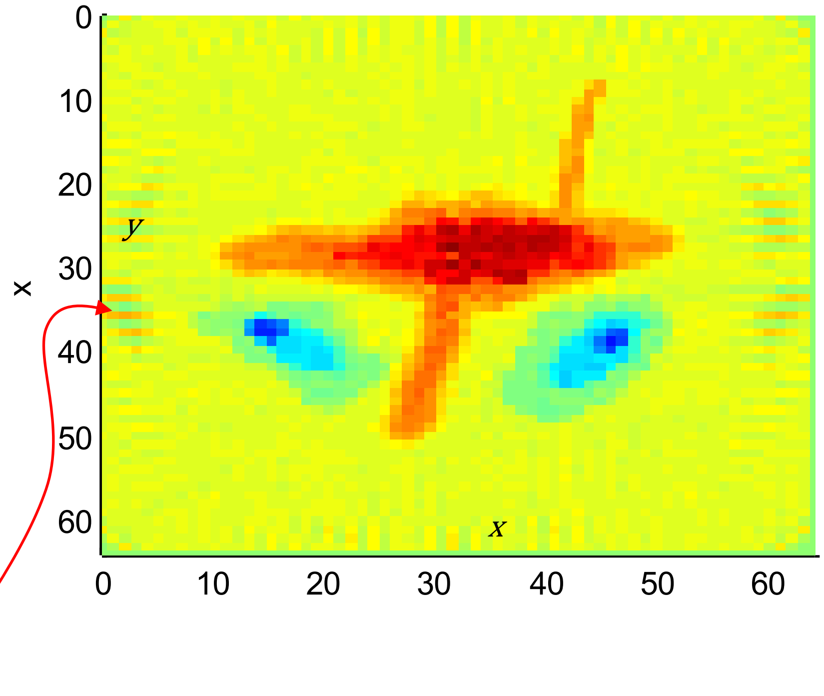
Estimated model



Estimated model



Estimated model

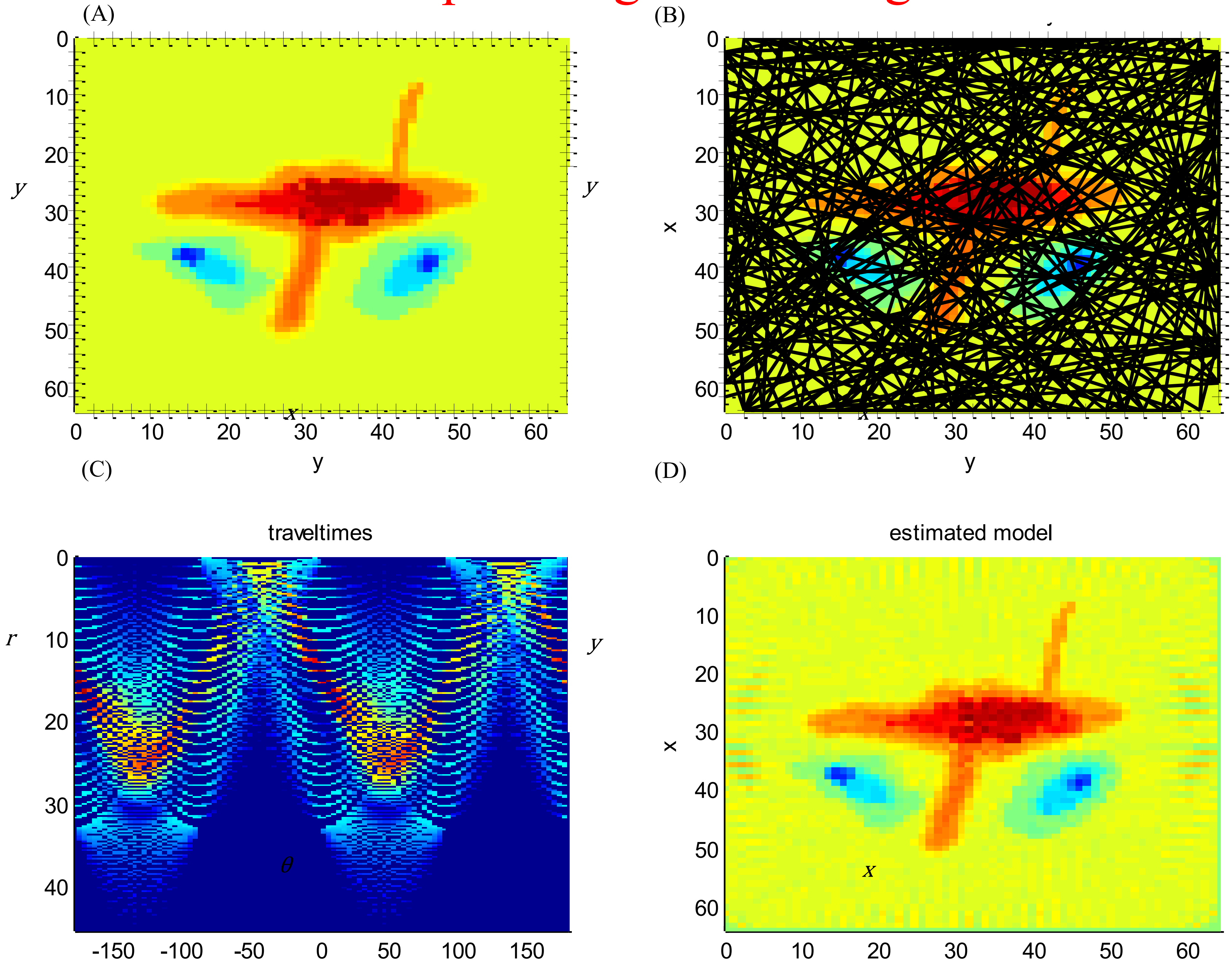


streaks due to minor data gaps

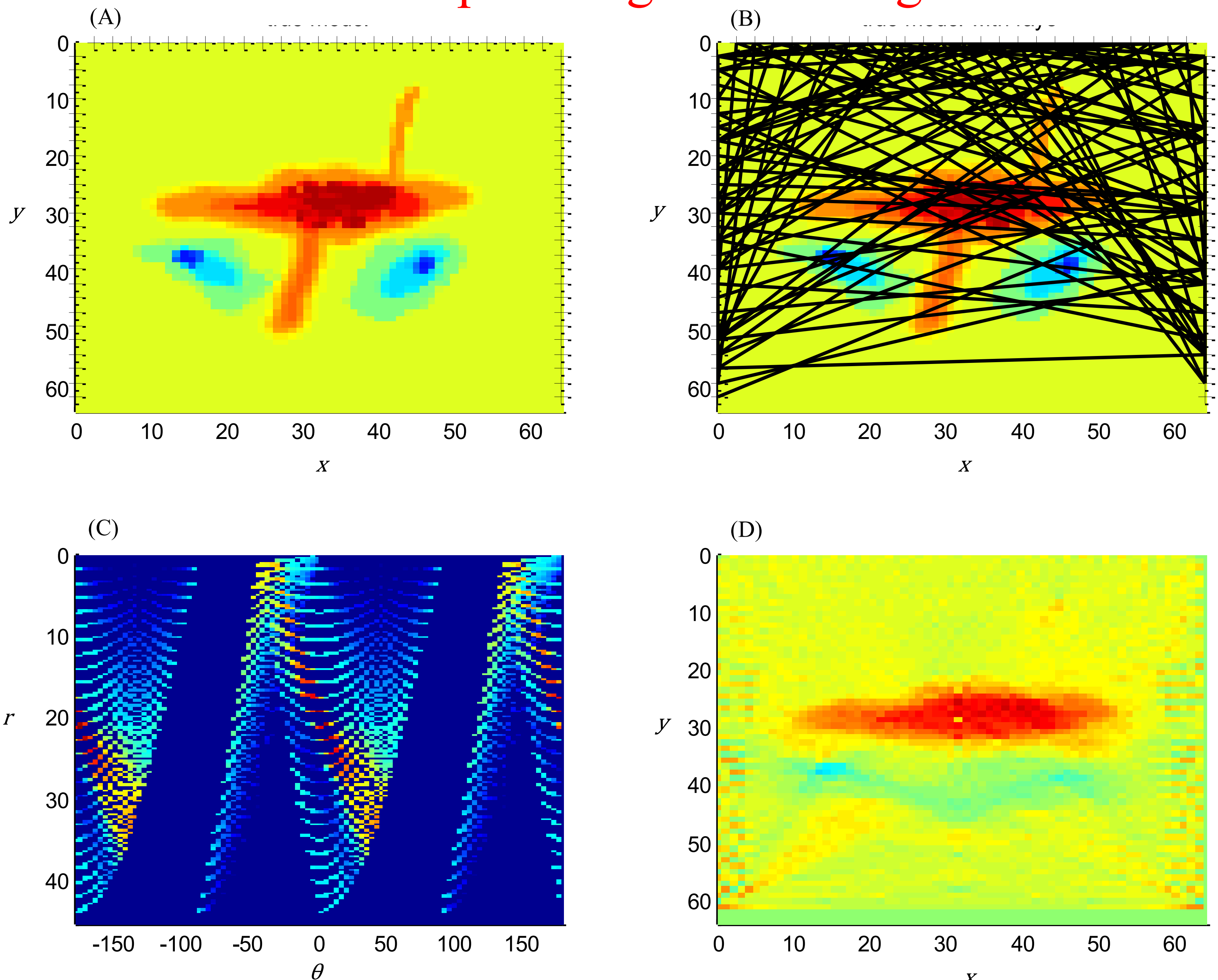
they disappear if ray density is doubled but what if the observational geometry is poor

so that broads swaths of rays are missing ?

complete angular coverage



incomplete angular coverage

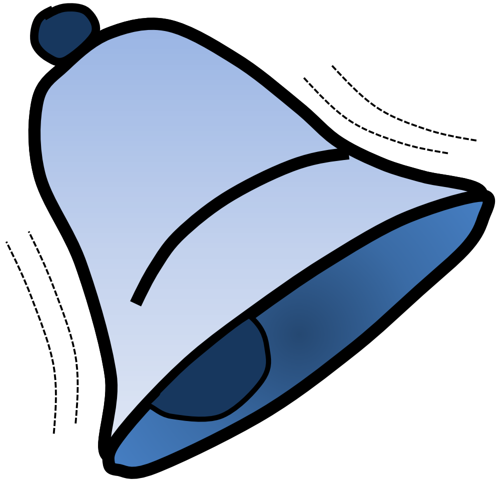


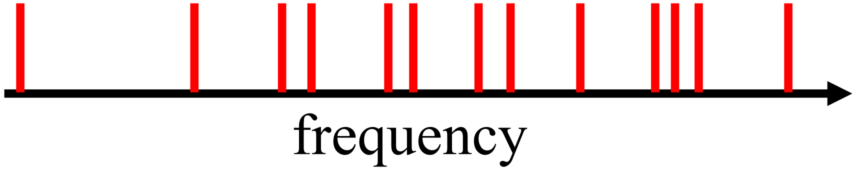
Part 2

vibrational problems

statement of the problem

Can you determine the structure of an object
just knowing the
characteristic frequencies at which it vibrates?





the Fréchet derivative
of frequency with respect to velocity
is usually computed using *perturbation theory*

hence a quick discussion of what that is ...

perturbation theory

a technique for computing an approximate solution to a complicated problem, when

1. The complicated problem is related to a simple problem by a small perturbation

2. The solution of the simple problem must be known

simple example

perturbation theory for a quadratic equation with a small first order term

$$x^2 + \varepsilon b x - c^2 = 0$$

assume ε is small and write solution as a power series in ε

$$x = x_0 + \varepsilon x_1 + \varepsilon^2 x_2 + \ \ldots$$

where x_0 solves the equation when $\varepsilon = 0$

$$x_0^2 - c^2 = 0$$

perturbation theory for a quadratic equation with a small first order term

$$x^2 + \varepsilon b x - c^2 = 0$$

assume ε is small and write solution as a power series in ε

$$x = x_0 + \varepsilon x_1 + \varepsilon^2 x_2 + \ \ldots$$

where x_0 solves the equation when $\varepsilon=0$

$$x_0^2 - c^2 = 0$$

we know the solution to this equation: $x_0 = \pm c$

plug the power series into the original equation

$$x^2 + \varepsilon b x - c^2 = 0$$

$$(x_0 + \varepsilon x_1 + \cdots)^2 + \varepsilon b(x_0 + \varepsilon x_1 + \cdots) - c^2 = 0$$

multiply out

$$x_0^2 + 2\varepsilon x_1 x_0 + \varepsilon b x_0 - c^2 + O(\varepsilon^2) = 0$$

Group terms of equal power in ε

$$(x_0^2 - c^2)\varepsilon^0 + (2x_1 x_0 + b x_0)\varepsilon + O(\varepsilon^2) = 0$$

since ε is arbitrary, coefficients of each power of ε must be individually zero

$$\varepsilon^0: \; (x_0^2 - c^2) = 0 \;\; \rightarrow \;\; x_0 = \pm c$$

$$\varepsilon^1: \; (2x_1 x_0 + b x_0) = 0 \;\; \rightarrow \;\; x_1 = -b/2$$

so solution is approximately

$$x = x_0 + \varepsilon x_1 + \; O(\varepsilon^2) = \; \pm c - \frac{\varepsilon b}{2} + \; O(\varepsilon^2)$$

since ε is arbitrary, coefficients of each power of ε must be individually zero

$$\varepsilon^0: \;\; (x_0^2 - c^2) = 0 \;\; \rightarrow \;\; x_0 = \pm c$$

$$\varepsilon^1: \;\; (2x_1 x_0 + b x_0) = 0 \;\; \rightarrow \;\; x_1 = -b/2$$

so solution is approximately

$$x = x_0 + \varepsilon x_1 + O(\varepsilon^2) = \pm c - \frac{\varepsilon b}{2} + O(\varepsilon^2)$$

note this agrees with the exact result computed from the quadratic formula

$$x = -\frac{\varepsilon b}{2} \pm \frac{\sqrt{\varepsilon^2 b^2 + 4c^2}}{2} = \pm c - \frac{\varepsilon b}{2} + O(\varepsilon^2)$$

Here's the actual vibrational problem

acoustic equation with
spatially variable sound velocity *v*

$$-\omega_n^2 \, p_n(\mathbf{x}) = v^2(\mathbf{x}) \, \nabla^2 p_n(\mathbf{x})$$

acoustic equation with spatially variable sound velocity v

$$-\omega_n^2\, p_n(\mathbf{x}) = v^2(\mathbf{x})\, \nabla^2 p_n(\mathbf{x})$$

frequencies of vibration
or
eigenfrequencies

patterns of vibration
or
eigenfunctions
or
modes

assume velocity can be written as a perturbation around some simple structure $v^{(0)}(\mathbf{x})$

$$v(\mathbf{x}) = v^{(0)}(\mathbf{x}) + \varepsilon v^{(1)}(\mathbf{x}) + \ldots$$

eigenfunctions known to obey orthonormality relationship

$$\int p_n(\mathbf{x})\, p_m(\mathbf{x})\, v^{-2}(\mathbf{x}) \mathrm{d}^3 x = \delta_{nm}$$

now represent eigenfrequencies and eigenfunctions as power series in ε

$$\omega_n = \omega_n^{(0)} + \varepsilon\omega_n^{(1)} + \varepsilon^2\omega_n^{(2)} + \cdots$$

$$p_n(\mathbf{x}) = p_n^{(0)}(\mathbf{x}) + \varepsilon p_n^{(1)}(\mathbf{x}) + \varepsilon^2 p_n^{(2)}(\mathbf{x}) + \cdots$$

now represent eigenfrequencies and eigenfunctions as power series in ε

$$\omega_n = \omega_n^{(0)} + \varepsilon\omega_n^{(1)} + \varepsilon^2\omega_n^{(2)} + \cdots$$

$$p_n(\mathbf{x}) = p_n^{(0)}(\mathbf{x}) + \varepsilon p_n^{(1)}(\mathbf{x}) + \varepsilon^2 p_n^{(2)}(\mathbf{x}) + \cdots$$

represent first-order perturbed shapes as sum of unperturbed shapes

$$p_m^{(1)} = \sum_{\substack{m \\ \omega_m \neq \omega_n}}^{\infty} b_{nm}\, p_m^{(0)}$$

plug series into original differential equation

group terms of equal power of ε

solve for first-order perturbation in eigenfrequencies $\omega_n^{(1)}$ and eigenfunction coefficients b_{nm}

(use orthonormality in process)

result

$$\omega_n^{(1)} = \omega_n^{(0)} \int \left[p_n^{(0)}(\mathbf{x})\right]^2 \left[v^{(0)}(\mathbf{x})\right]^{-3} v^{(1)}(\mathbf{x})\, \mathrm{d}^3x$$

$$b_{nm} = \frac{2\left(\omega_m^{(0)}\right)^2}{\left(\omega_m^{(0)}\right)^2 - \left(\omega_n^{(0)}\right)^2} \int p_n^{(0)}(\mathbf{x})\, p_m^{(0)}(\mathbf{x}) \left[v^{(0)}(\mathbf{x})\right]^{-3} v^{(1)}(\mathbf{x})\, \mathrm{d}^3x$$

result for eigenfrequencies

$$\omega_n^{(1)} = \omega_n^{(0)} \int \left[p_n^{(0)}(\mathbf{x})\right]^2 \left[v^{(0)}(\mathbf{x})\right]^{-3} v^{(1)}(\mathbf{x})\, \mathrm{d}^3x$$

write as standard inverse problem

$$\omega_n^{(1)} = \int G_n(\mathbf{x})\, v^{(1)}(\mathbf{x})\, \mathrm{d}^3x \quad \text{with} \quad G_n(\mathbf{x}) = \omega_n^{(0)}\left[p_n^{(0)}(\mathbf{x})\right]^2\left[v^{(0)}(\mathbf{x})\right]^{-3}$$

standard continuous inverse problem

$$\omega_n^{(1)} = \int G_n(\mathbf{x})\, v^{(1)}(\mathbf{x})\, \mathrm{d}^3x \quad \text{with} \quad G_n(\mathbf{x}) = \omega_n^{(0)} \left[p_n^{(0)}(\mathbf{x})\right]^2 \left[v^{(0)}(\mathbf{x})\right]^{-3}$$

standard continuous inverse problem

$$\omega_n^{(1)} = \int G_n(\mathbf{x})\, v^{(1)}(\mathbf{x})\, \mathrm{d}^3x \quad \text{with} \quad G_n(\mathbf{x}) = \omega_n^{(0)}\left[p_n^{(0)}(\mathbf{x})\right]^2\left[v^{(0)}(\mathbf{x})\right]^{-3}$$

perturbation in the eigenfrequencies are the data

perturbation in the velocity structure is the model function

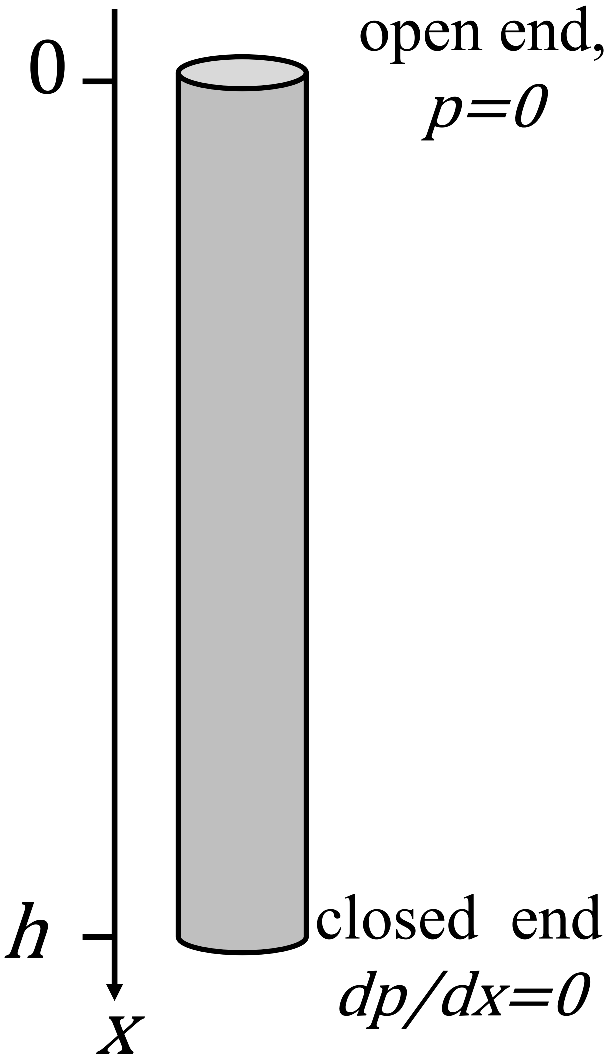
standard continuous inverse problem

data kernel or Fréchet derivative

$$\omega_n^{(1)} = \int G_n(\mathbf{x})\, v^{(1)}(\mathbf{x})\, \mathrm{d}^3x \quad \text{with} \quad G_n(\mathbf{x}) = \omega_n^{(0)} \left[p_n^{(0)}(\mathbf{x})\right]^2 \left[v^{(0)}(\mathbf{x})\right]^{-3}$$

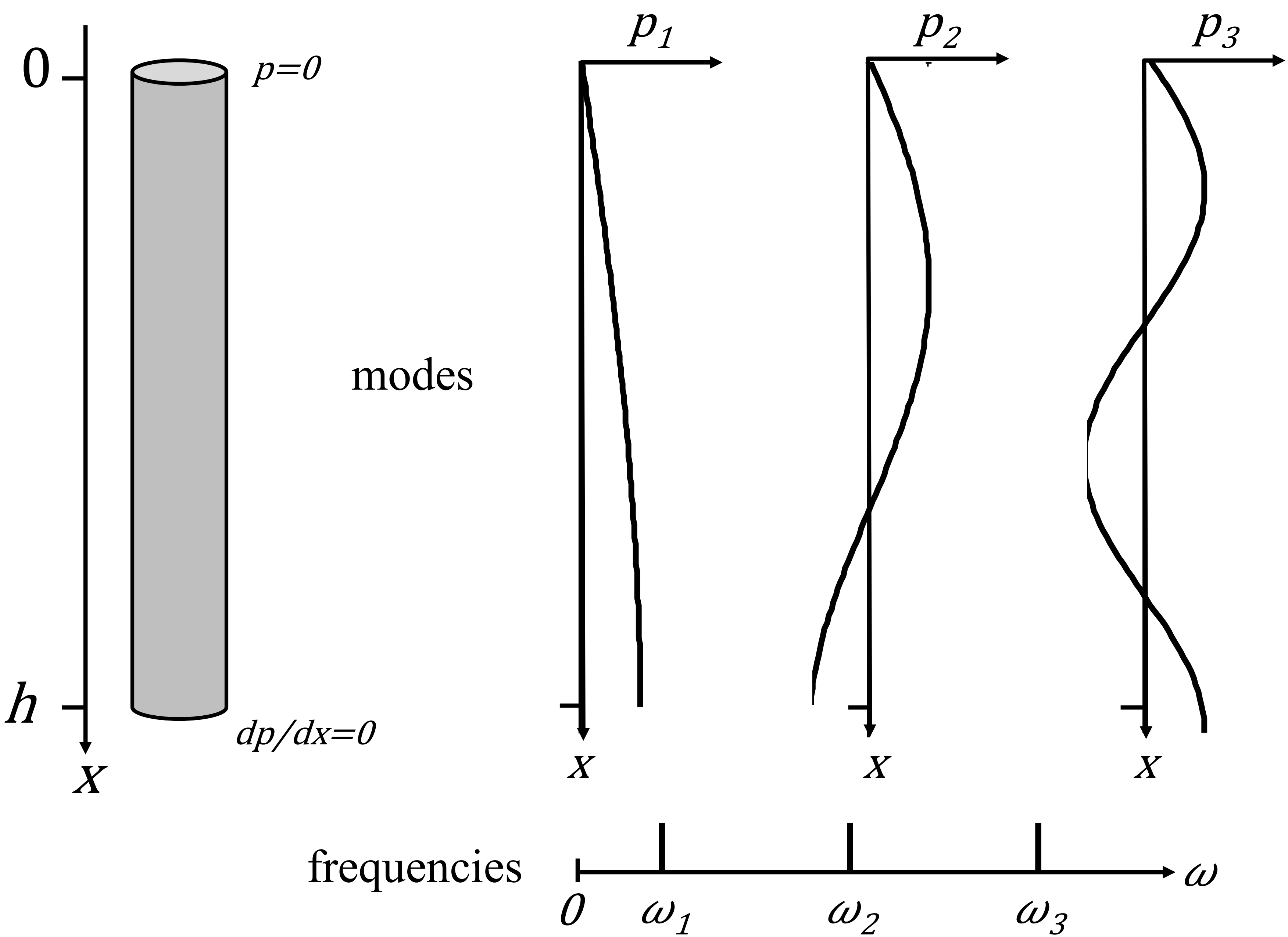
depends upon the
unperturbed velocity structure,
the unperturbed eigenfrequency
and the unperturbed mode

1D organ pipe



unperturbed problem has constant velocity

perturbed problem has variable velocity



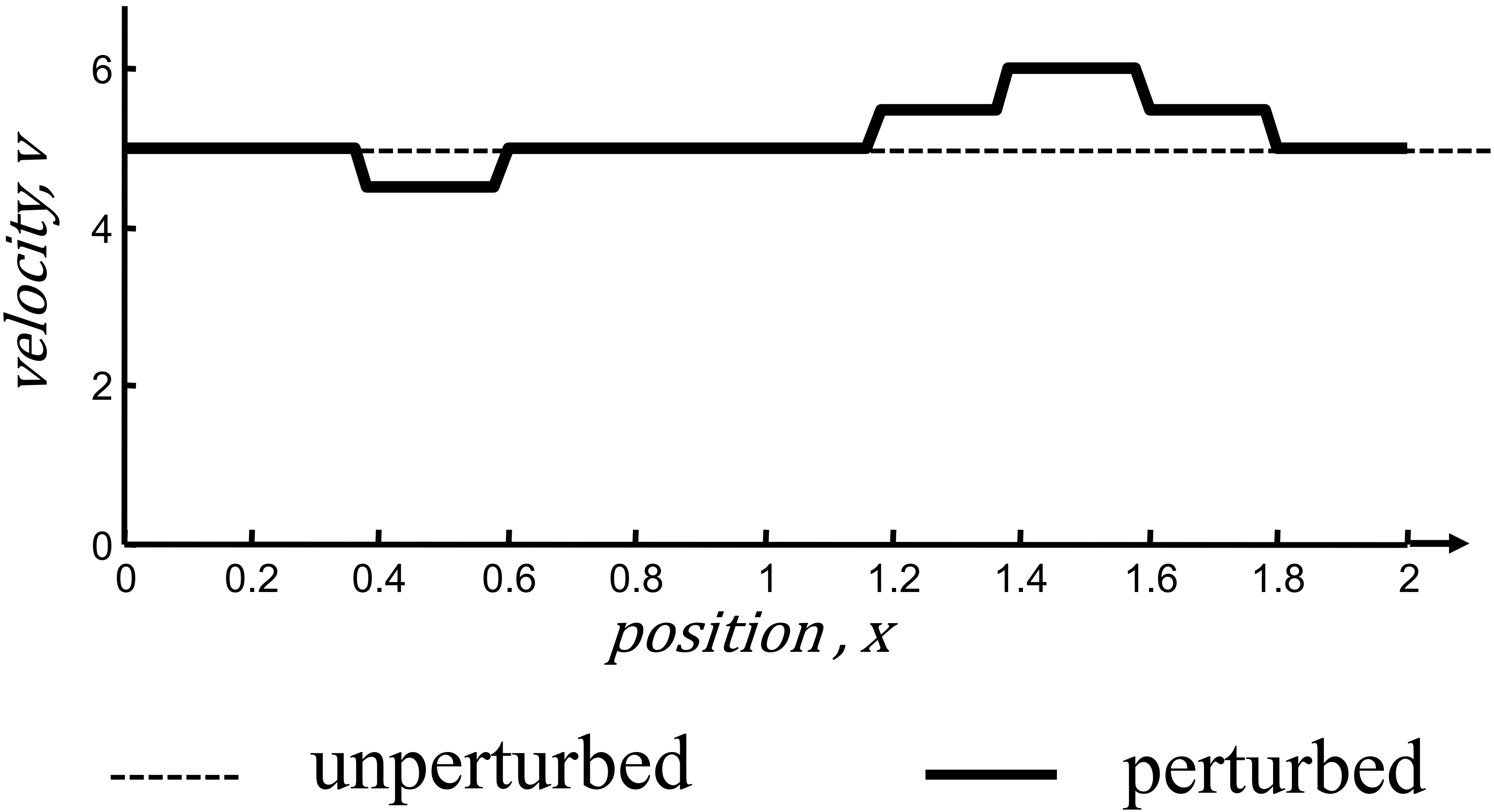
0
p=0
h
dp/dx=0
x
modes
p_1
p_2
p_3
x
x
x
frequencies
0
ω_1
ω_2
ω_3
ω

solution to unperturbed problem

$$p_n^{(0)}(x) = \frac{2\left[v^{(0)}\right]^2}{h} \sin\left\{\frac{(n - ½)\pi}{h} x\right\}$$

$$\omega_n^{(0)} = \frac{\pi(n - ½)v^{(0)}}{h} \quad \text{with } n = 1,2,3,\cdots$$

velocity structure

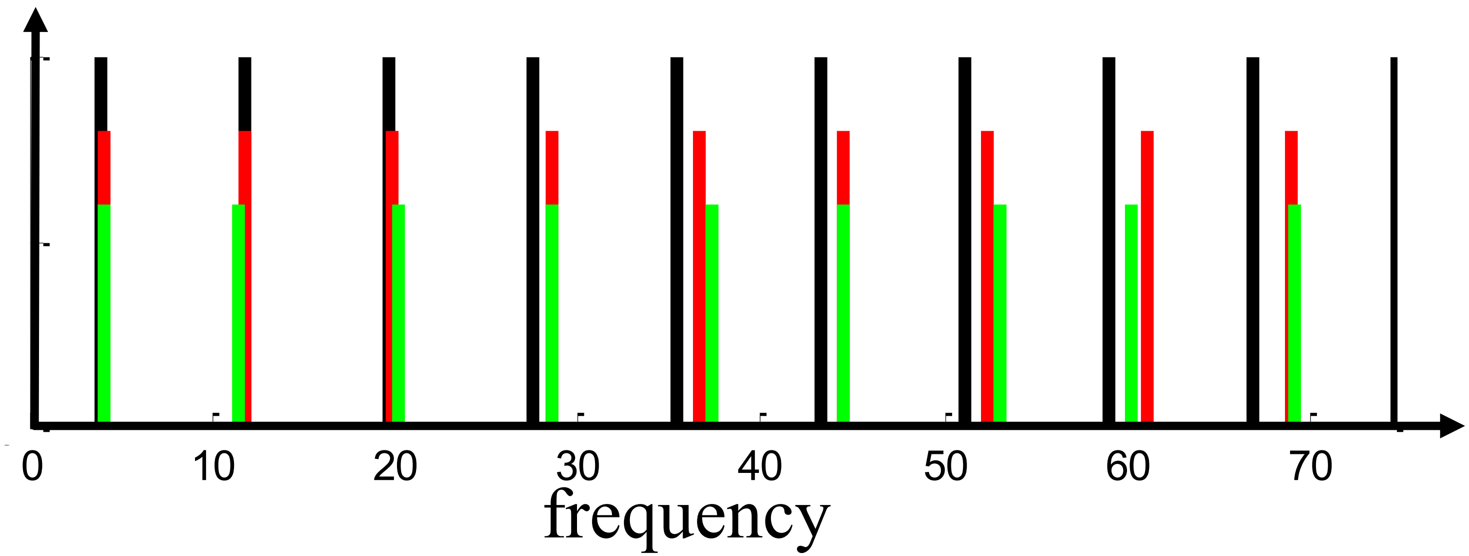


How to discretize the model function?

our choice is very simple

m is veloctity function evaluated at sequence of points equally spaced in x

the data
a list of frequencies of vibration

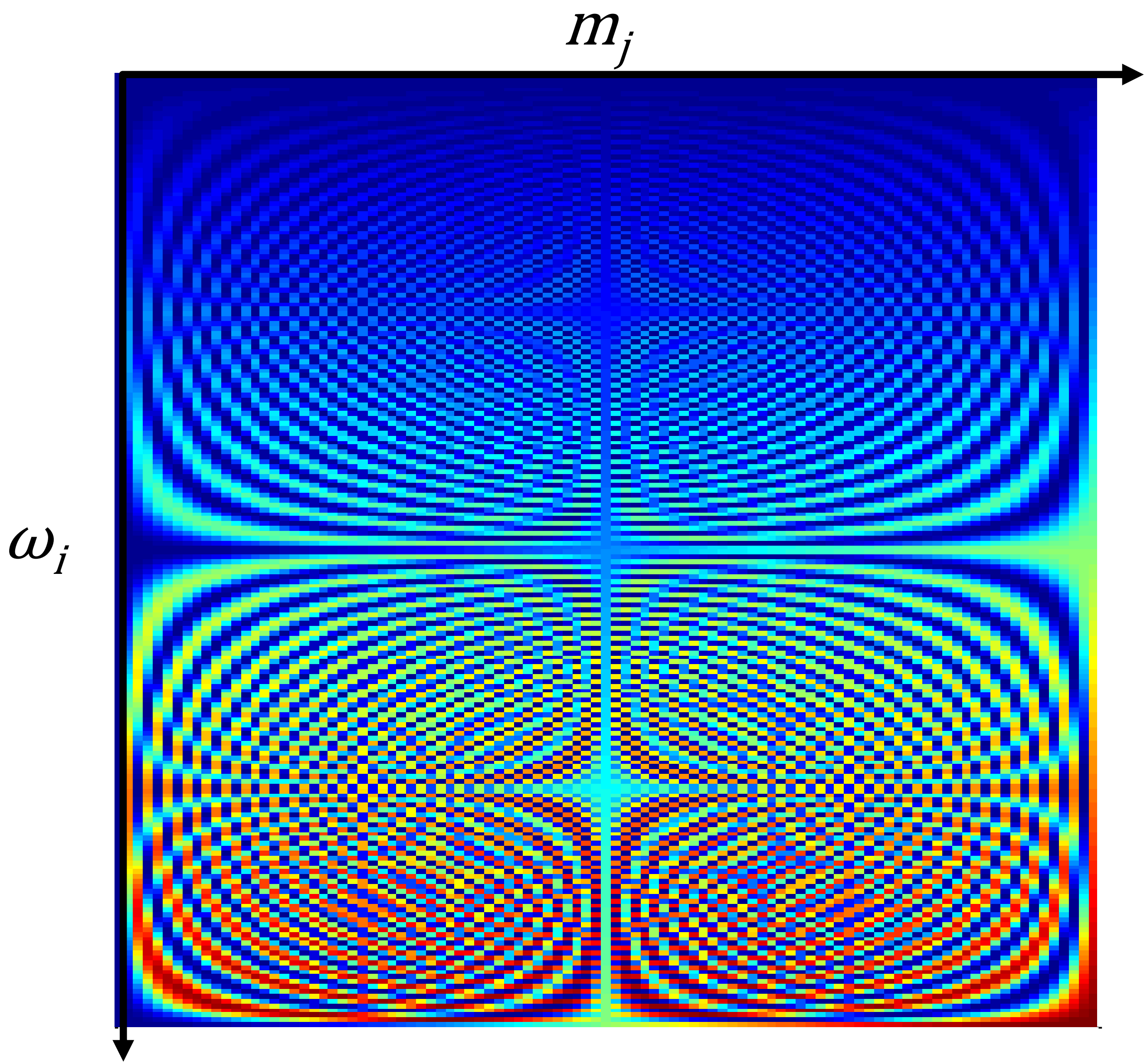


true, unperturbed

true, perturbed

observed = true, perturbed + noise

the data kernel



Solution Possibilities

1. Damped Least Squares (implements smallness):

   Matrix **G** is not sparse

   use `bicg()` with damped least squares function

2. Weighted Least Squares (implements smoothness):

   Matrix **F** consists of **G** plus

   second derivative smoothing

   use `bicg()` with weighted least squares function

Solution Possibilities

1. Damped Least Squares (implements smallness):

   Matrix **G** is not sparse

   use **`bicg()`** with damped least squares function

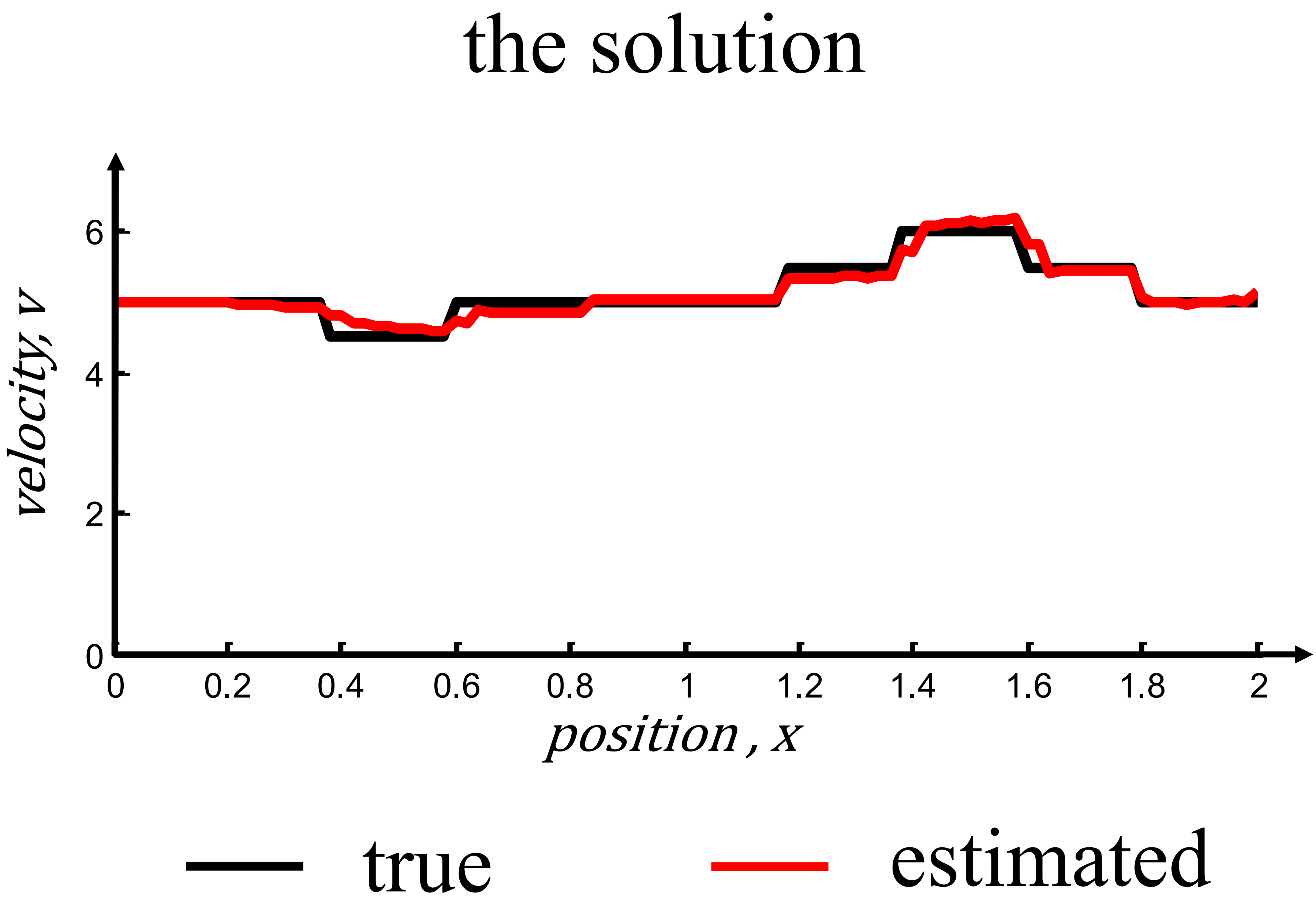
   our choice

2. Weighted Least Squares (implements smoothness):

   Matrix **F** consists of **G** plus

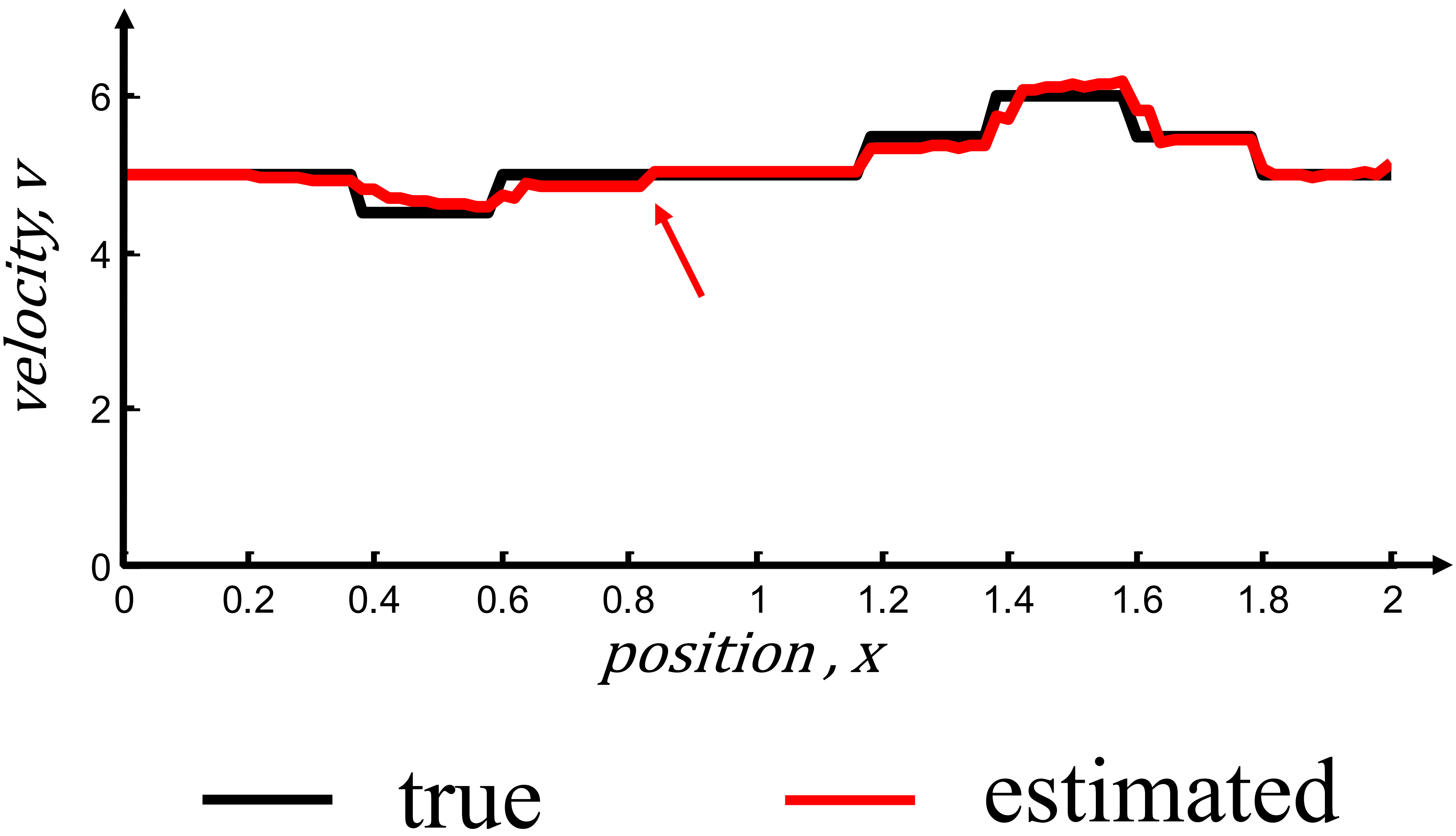
   second derivative smoothing

   use **`bicg()`** with weighted least squares function

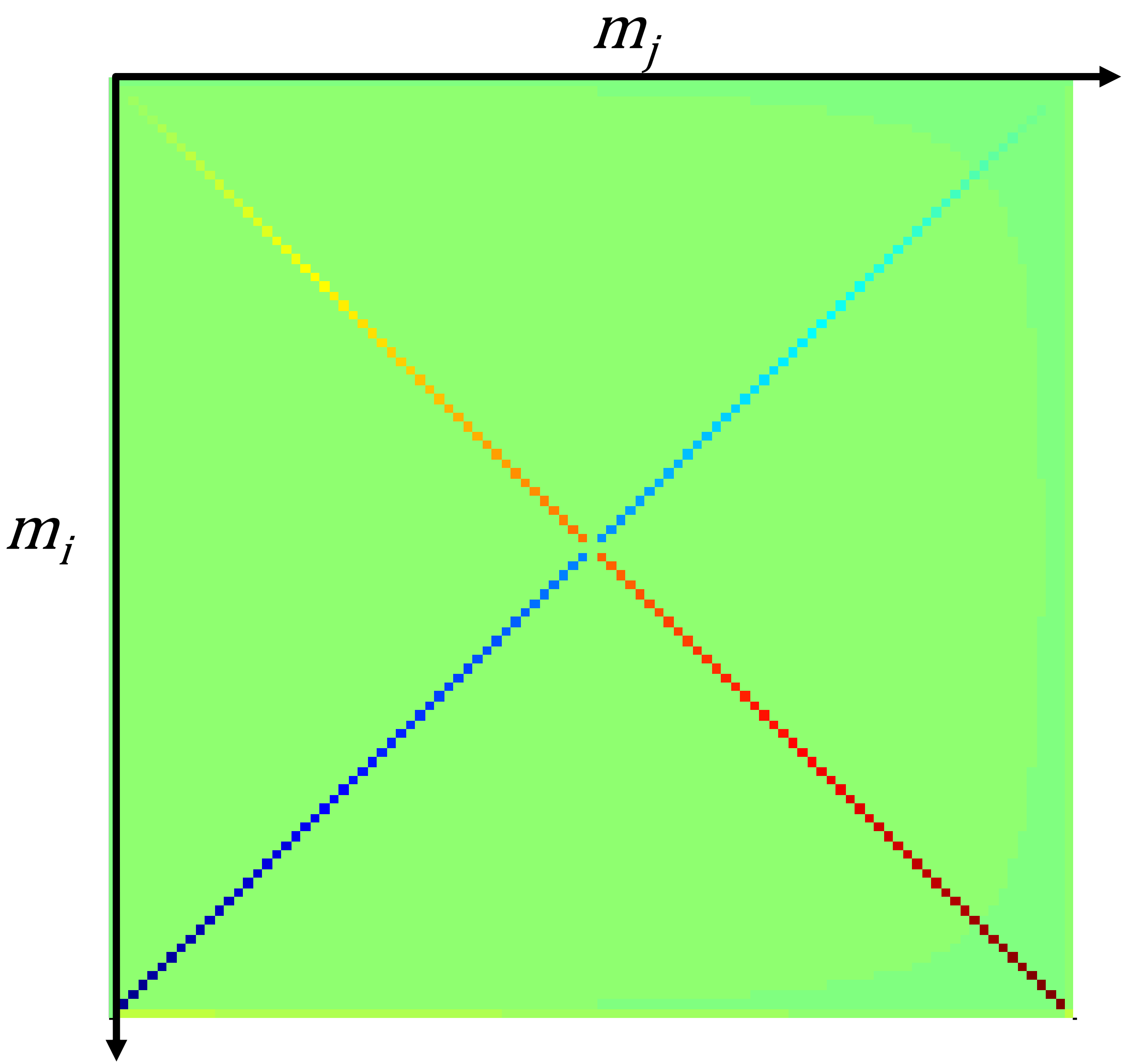


the solution
velocity, v
6
4
2
0
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
position , x
true
estimated

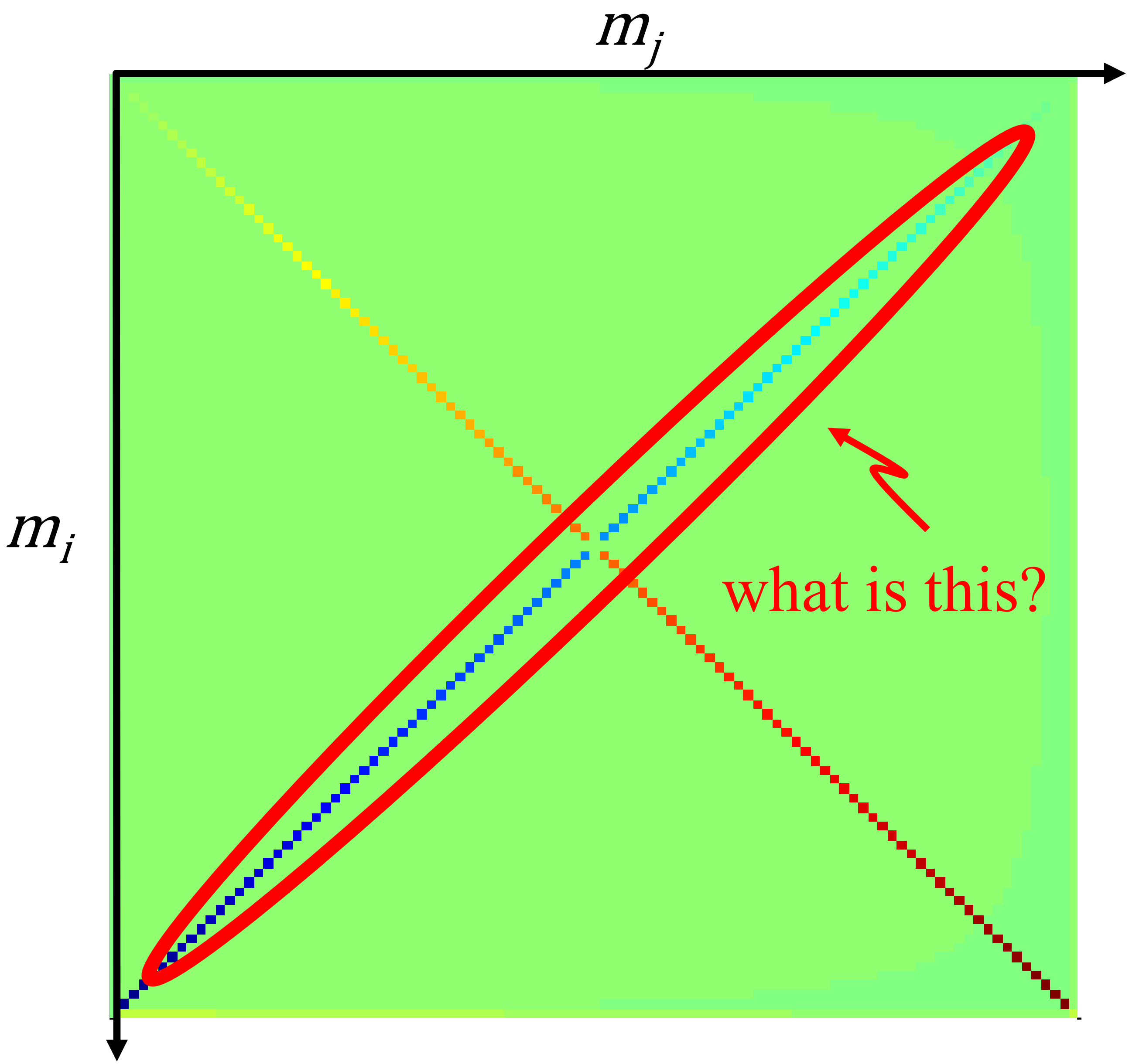
the solution



the model resolution matrix



the model resolution matrix



This problem has a type of nonuniqueness

that arises from its symmetry

a positive velocity anomaly at one end of the organ pipe

trades off with a negative anomaly at the other end

this behavior is very common
and is why eigenfrequency data
are usually supplemented with other data

e.g. travel times along rays

that are not subject to this nonuniqueness

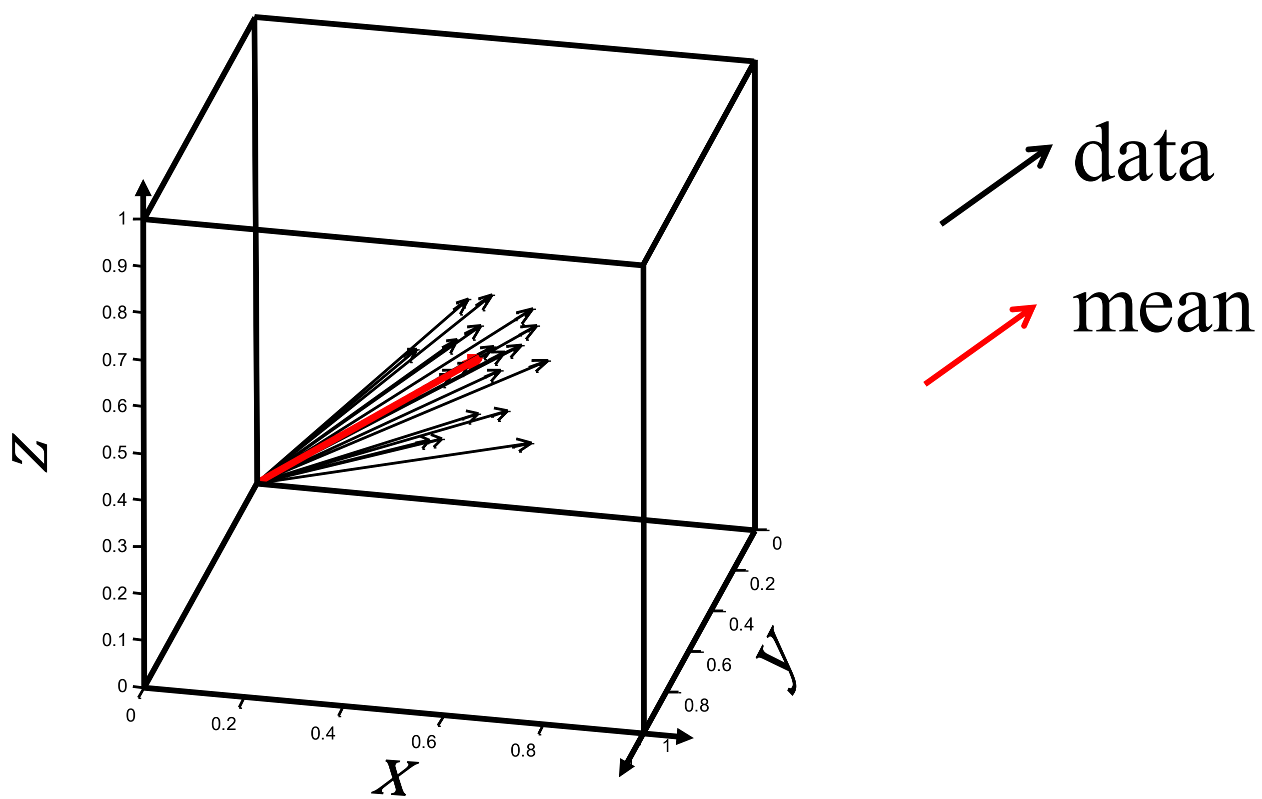
Part 3

determining mean directions

statement of the problem

you measure a bunch of directions (unit vectors)

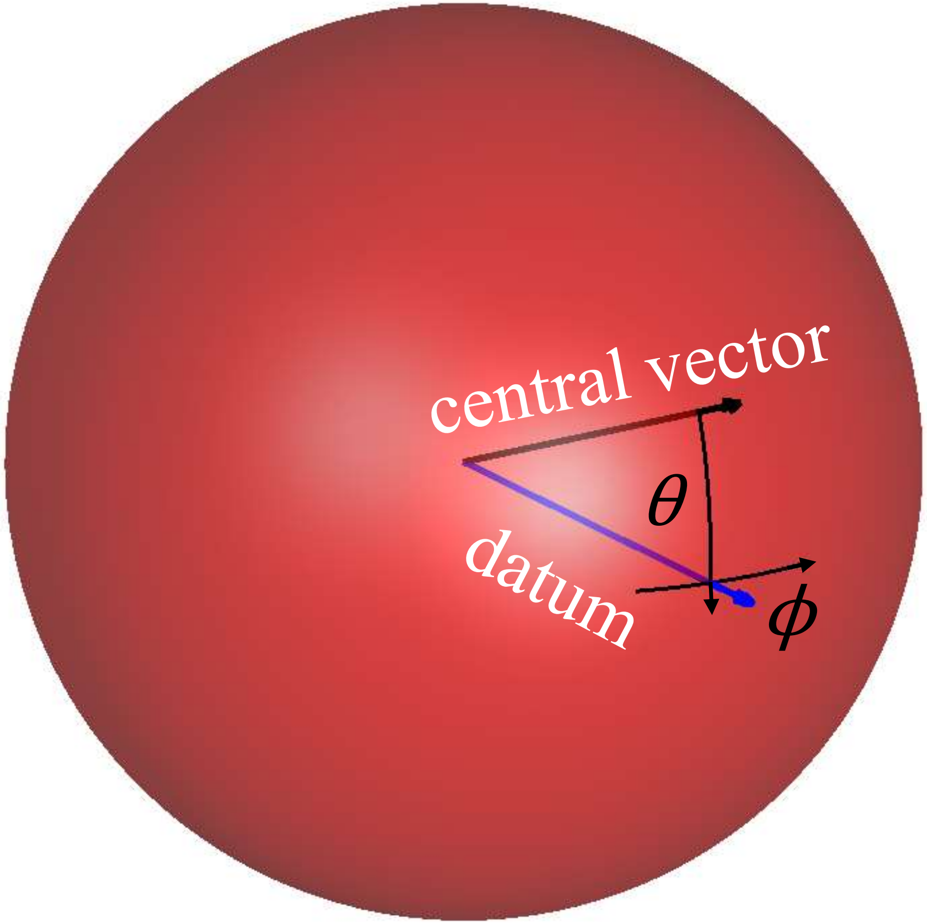
what's their mean?



what's a reasonable probability density function for directional data?

Gaussian doesn't quite work
because
its defined on the wrong interval
$(-\infty, +\infty)$

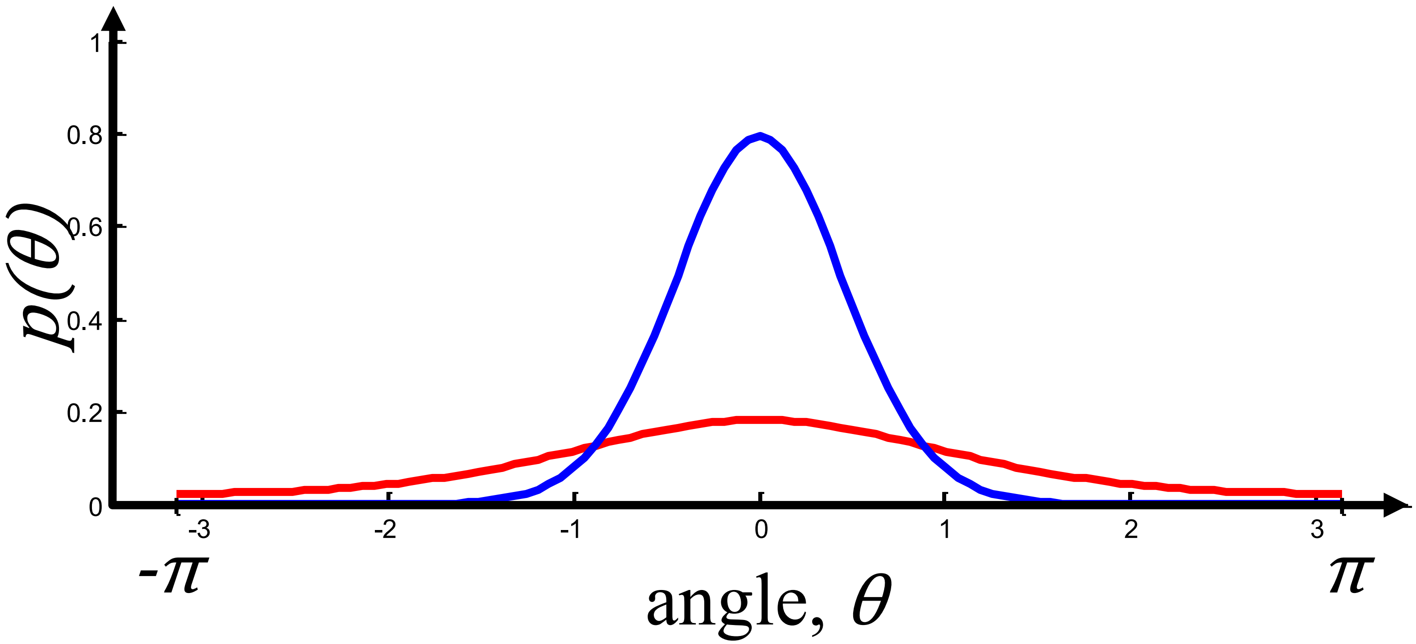
coordinate system



distribution should be symmetric in ϕ

Fisher distribution

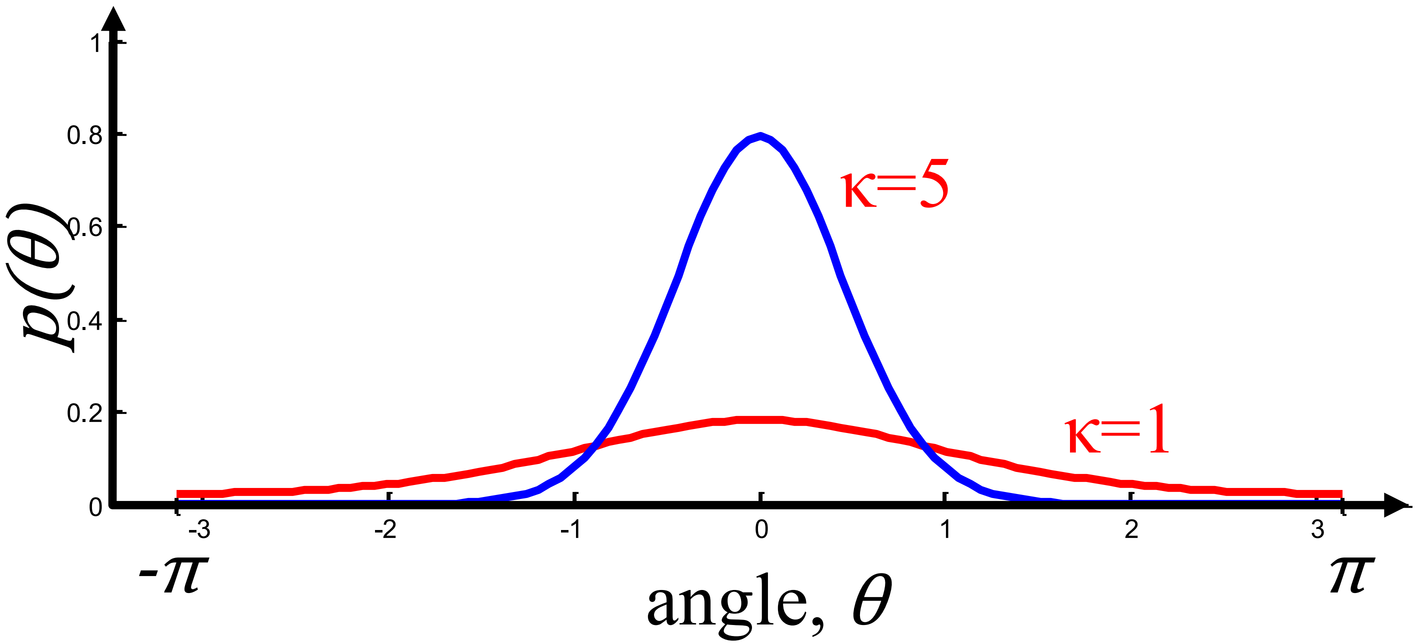
similar in shape to a Gaussian but on a sphere



$$p(\theta, \phi) = \frac{\kappa}{4\pi \sinh(\kappa)} \exp[\kappa \cos(\theta)]$$

Fisher distribution

similar in shape to a Gaussian but on a sphere



"precision parameter" quantifies width of p.d.f.

$$p(\theta, \phi) = \frac{\kappa}{4\pi \sinh(\kappa)} \exp[\kappa \cos(\theta)]$$

solve by

direct application of

principle of maximum likelihood

maximize joint p.d.f. of data

$$p(\theta, \phi) = \left[\frac{\kappa}{4\pi \sinh(\kappa)}\right]^N \exp\left[\kappa \sum_{i=1}^{N} \cos(\theta_i)\right] \prod_{i=1}^{N} \sin(\theta_i)$$

with respect to
κ and cos(θ)

x: Cartesian components of observed unit vectors

m: Cartesian components of central unit vector; must constrain $|\mathbf{m}|=1$

$$\cos(\theta_i) = \mathbf{x}^{\mathrm{T}}\mathbf{m} = [x_i m_1 + y_i m_2 + z_i m_3]$$

likelihood function

$$L = \log(p) = N\log(\kappa) - \mathrm{N}\log(4\pi) - N\log[\sinh(\kappa)]$$

$$+ \kappa \sum_{i=1}^{N} [x_i m_1 + y_i m_2 + z_i m_3] + \sum_{i=1}^{N} \log[\sin(\theta_i)]$$

constraint

$$C = \sum_i m_i^2 - 1 = 0$$

unknowns

m, κ

Lagrange multiplier equations

$$\kappa \sum_i x_i - 2\lambda m_1 = 0$$

$$\kappa \sum_i y_i - 2\lambda m_2 = 0$$

$$\kappa \sum_i z_i - 2\lambda m_3 = 0$$

$$\frac{N}{\kappa} - N\frac{\cosh(\kappa)}{\sinh(\kappa)} + \sum_{i=1}^{N}[x_i m_1 + y_i m_2 + z_i m_3] = 0$$

Results

$$[m_1, m_2, m_3]^{\mathrm{T}} = \frac{[\sum_i x_i, \sum_i y_i, \sum_i z_i]^{\mathrm{T}}}{\{(\sum_i x_i)^2 + (\sum_i y_i)^2 + (\sum_i z_i)^2\}^{1/2}}$$

$$\kappa \approx \frac{\mathrm{N}}{\mathrm{N} - \sum_{\mathrm{i}} \cos(\theta_i)}$$

valid when $\kappa > 5$

Results

$$[m_1, m_2, m_3]^{\mathrm{T}} = \frac{[\sum_i x_i, \sum_i y_i, \sum_i z_i]^{\mathrm{T}}}{\{(\sum_i x_i)^2 + (\sum_i y_i)^2 + (\sum_i z_i)^2\}^{1/2}}$$

central vector is parallel to the vector that you get by putting all the observed unit vectors end-to-end

Solution Possibilities

Determine **m** by evaluating simple formula

1. Determine κ using simple but approximate formula

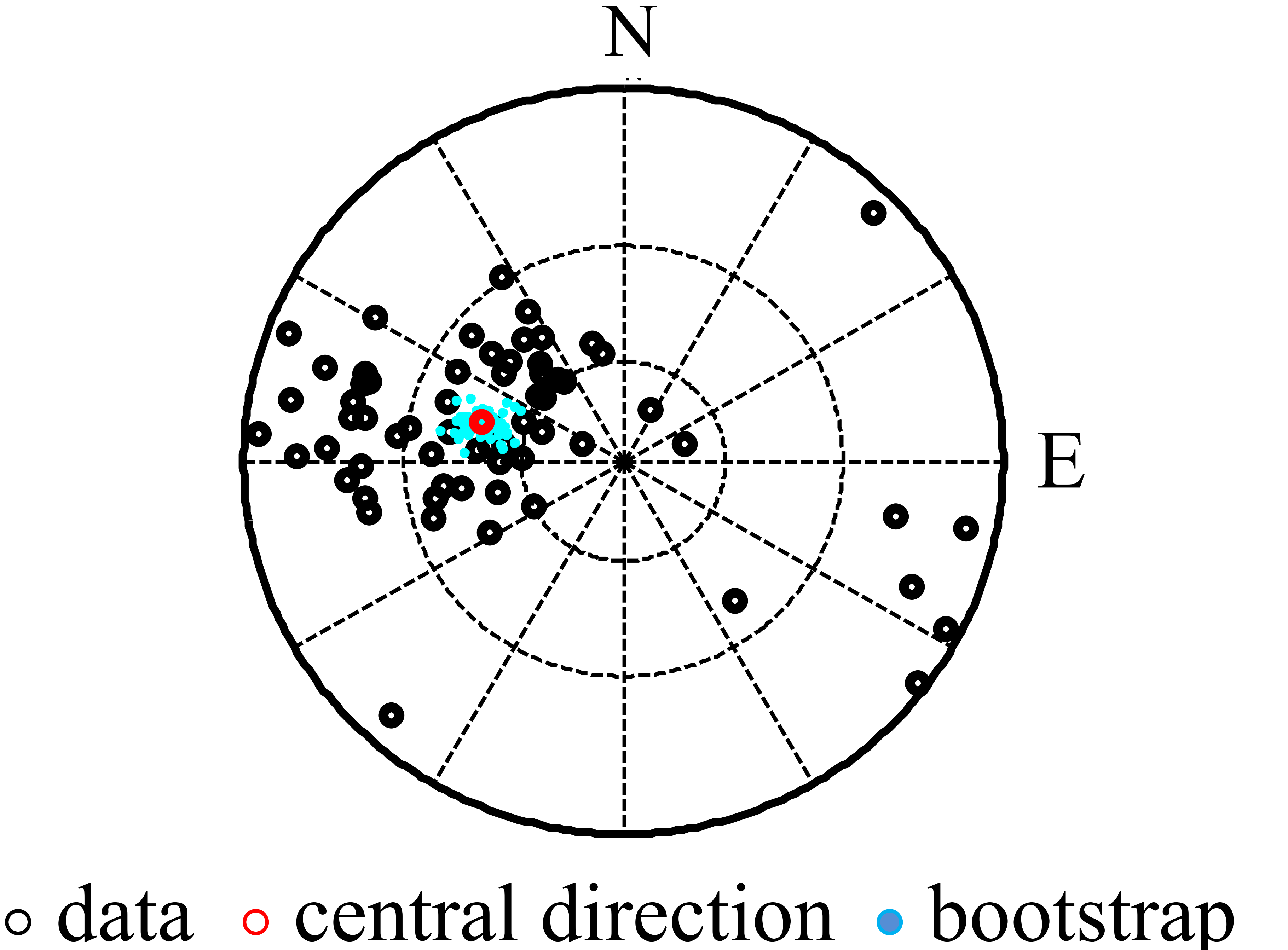
only valid when $\kappa > 5$

2. Determine κ using bootstrap method

our choice

Application to Subduction Zone Stresses

Determine the mean direction of
P-axes
of deep (300-600 km) earthquakes
in the Kurile-Kamchatka subduction zone


N
E
data
central direction
bootstrap