INTERNATIONAL ATOMIC ENERGY AGENCY
UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION

INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS
34100 TRIESTE (ITALY) - P.O.B. 586 - MIRAMARE - STRADA COSTIERA 11 - TELEPHONE: 2240-1
CABLE: CENTRATOM - TELEX 460892-I

H4.SMR/179-14

AUTUMN COURSE ON SEISMOLOGY

(1 September - 3 October 1986)

INVERSE PROBLEMS - II

B.A. HOBBS

University of Edinburgh
Edinburgh, U.K.

1. Inverse Theory

1.1 Introduction

A geophysicist wants to learn something about the earth. In order to study a particular aspect, or parameter, of the earth some measurements must be made that relate to that parameter. For example suppose we are interested in the earth's density. There are several measurable quantities that depend on the density distribution within the earth. If we determine the times of travel of seismic waves from a particular source to a number of receivers around the world, we can then determine the variation of seismic wave velocities with depth in the earth. The Adams-Williamson equation then yields the density distribution. Alternatively, after a large earthquake, the normal modes of vibration of the earth are exited and the earth may continue vibrating for many days. The period of each normal mode depends on the distibution of density, bulk modulus and shear modulus in the earth. As another example suppose we are interested in the earth's conductivity distribution. The response of the earth to a periodic magnetic field variation of a given frequency is a measure of the electric currents induced in the earth and these depend on the distibution of conductivity. There are many such examples in geophysics. The inverse problem is simply, given the measurements, find the given parameter.

Clearly there is only ever a finite amount of data for the investigation of any given problem. For example the periods of several hundred normal modes of vibration are known - obtained from Fourier transforms of seismogram records. The exitation of higher frequencies have small amplitudes and are

lost in the instrumental noise level. In the conductivity problem, theoretical considerations limit knowledge of the response at very high and very low frequencies. Also it is not possible to obtain an infinite number of Fourier transforms from data for in performing the Fourier transform the record has to be sampled at small intervals. The choice of these intervals limits the number of independent transforms. Clearly then data are finite in number - this has important consequences, it means that any earth model found is non-unique.

The starting point then is a set of measurements

$$e_j \quad j=1,2,\ldots,N \tag{1}$$

and these are necessarily finite in number. How are these numbers to be interpreted? First, methods may be classified into "forward" and "inverse".

1.1.1 Forward methods

Let r represent the variation of radius within the earth and let m(r) be a model - in other words a distribution of the parameter of interest. Calculate the behaviour, or response, of the model $F_j(m)$ and compare this with the data e_j. Monte-Carlo methods come into this category whereby a statistically random search of models is made and only those whose responses fit the data to within a prescribed accuracy are retained.

1.1.2 Inverse methods

Given the behaviour of the earth (namely the data e_j), calculate the model m(r).

There are many categories of inverse problems. One can approach the subject through continuous or discrete theory, there are over-determined, exact and underdetermined cases

and the problem can be linear or non-linear. I will illustrate the concept of inverse theory through a developement of continuous theory.

1.2 Linear problems

It is frequently possible to relate a model m(r) to the data e_i by a linear relationship (a Fredholm integral equation of the first kind) of the form

$$e_i = \int_0^1 G_i(r)\, m(r)\, dr \qquad (2)$$

For example the mass of the earth can be related to its density by

$$\text{mass} = \int_0^1 4\pi r^2 \rho(r)\, dr \qquad (3)$$

and the moment of inertia by

$$\text{M.I.} = \int_0^1 (8\pi r^4/3)\, \rho(r)\, dr \qquad (4)$$

When this is not possible, we may have to resort to linearised inverse theory

1.3 Uniqueness

We have admitted to only having a finite number of data values for use in eq(2). Under these conditions Backus and Gilbert(1967) have proved that the class of models satisfying the data is either empty or infinite dimensional. One source of non-uniqueness is therefore associated with the data being

inadequate (ie finite in number). (Even if an infinite amount of data were available, solution to eq(2) might still be non-unique) The simplest of examples from discrete theory illustrates this point. Suppose the "model" m is represented by a straight line, requiring two parameters for its description. Suppose further that the data e_j comprises only one point (j=1). Then an infinite number of straight lines pass through that one point - even if the one point is accurately located

The second source of non-uniqueness stems from errors in the data values e_i - the data are thus inaccurate as well as being inadequate. In the above simple example, many more straight lines will pass through the area in which the datum is known to lie within some given probability

In the light of the above discussion there are two aspects of inverse theory to be addressed First, which model from the infinite dimensional space should be selected and how and second, how is this model to be interpreted - how certain can we be of the characteristics of the model? The model finding and model interpretation aspects will be considered using continuous inverse theory

2. The Backus-Gilbert formulation

Hilbert-space and Fréchet Differentiability

Let r be a real variable such that $0 \leq r \leq 1$ (in general r will be the normalised earth radius) and let m(r) be an earth model - that is m represents some distribution of parameters throughout the earth. As examples m may represent a density model $m(r) = \rho(r)$, or a conductivity model $m(r) = \sigma(r)$ or an n-tuple of the form $m(r) = (\rho(r), \kappa(r), \mu(r))$. The model m(r) must be real-valued, piecewise continuous. For any two earth models m(r) and m'(r), their linear combination is the earth model $\alpha m(r) + \alpha' m'(r)$ where α and α' are two real numbers. With this definition earth models form points in an infinite-dimensional linear space M, the space of all conceivable earth models. By defining an inner product and a norm, M becomes an inner product space which

can be completed to a Hilbert space. It is convenient to define the inner product as

$$(m,m') = \int_0^1 m(r)m'(r)\,dr \qquad (5)$$

and the norm as

$$\|m\| = (m,m)^{1/2} \qquad (6)$$

The data obtained, e_j ,j=1,2,...,N are related to the model m via

$$e_j = F_j(m) \qquad j=1,2,\ldots,N \qquad (7)$$

F is simply a rule which allows calculation of the response of the earth model m, i.e. F is a functional on the space M. The simplest functionals are linear functionals as in eq(2), where

$$F(m) = (G,m) = \int_0^1 G(r)m(r)\,dr \qquad (8)$$

with G independent of m.

The analysis may be extended to non-linear functionals by considering Fréchet differentiation, that is: how does a small change in the model affect the functional, or: what is the derivative of the functional with respect to the model? For linear functionals, from eq(8),

$$F(m+\delta m) = F(m) + \int_0^1 G(r)\delta m(r)\,dr \qquad (9)$$

that is

$$F(m+\delta m) = F(m) + (G,\delta m) \qquad (10)$$

If, however, we may write

$$F(m+\delta m) = F(m) + (G,\delta m) + O\|\delta m\|^2 \qquad (11)$$

where $D \in M$, then the datum to which F relates is Fréchet differentiable and G is the Fréchet derivative. Clearly linear functionals are Fréchet differentiable with Fréchet derivative G. For the examples above, mass is Fréchet differentiable with $G=4\pi r^2$, moment of inertia has $G=8\pi r^4/3$. A non-linear example is provided by $\tau(p)$ inversion where

$$\tau(p) = 2\int_0^{depth} \left(u^2(z) - p^2\right)^{1/2} dz \qquad (12)$$

Looking at small perturbations,

$$\delta\tau(p) = 2\int_0^{depth} u(z)\delta u(z)\left(u^2(z) - p^2\right)^{-1/2} dz + O\|\delta u\|^2 \qquad (13)$$

where the Fréchet derivative is

$$G(z) = u(z)\left(u^2(z) - p^2\right)^{-1/2} \qquad (14)$$

and this depends on the model u(z).

Thus to first order in m, eq(11) may be written

$$F(m+\delta m) = F(m) + (G,\delta m) \qquad (15)$$

This equation is exact when F is linear and is the linearised version (which may be approximate) otherwise. We are now in a position to find a model fitting observations.

2.1 Generating a model

Given a set of data

e_i, j=1,2.....,N (16)

we may propose a model m(r) and calculate the functionals $F_j(m)$ to which the e_j relate. Unless we are exceptionally good at guessing, it is likely that the calculated values $F_j(m)$ will not agree with the data e_j . How can we now change the model m(r) to m(r) + δm(r) in order to obtain agreement? Let us consider changing the model as little as possible in order that

$$F_j(m + \delta m) = e_j \qquad j=1,2,...N \tag{17}$$

One measure of the closeness of m to m + δm is the smallmess of the norm

$$\|\delta m\|^2 = (\delta m,\delta m) = \int_0^1 (\delta m)^2 \, dr \tag{18}$$

We ask then, what δm minimises (18) subject to the constraints (17)? Using eq(15) the constraints become

$$(G_j,\delta m) = e_j - F_j(m) \qquad j=1,2,....N \tag{19}$$

This is just a minimisation problem with N constraints. Introduce Lagrange multipliers β_j, j=1,2....,N and minimise

$$S = (\delta m,\delta m) + \sum_{j=1}^{N} \beta_j \left((G_j,\delta m) - e_j + F_j(m) \right) \tag{20}$$

It is actually more convenient to set $\beta_j = -2\alpha_j$, then minimising (20) yields

$$\delta m = \sum_{j=1}^{N} \alpha_j G_j \tag{21}$$

Substituting (21) into (19) the coefficients may be determined

from the linear equations

$$\sum_{k=1}^{N} \alpha_k (G_j,G_k) = e_j - F_j(m) \qquad j=1,2, ..N \tag{22}$$

More explicitly the equations are

$$\sum_{k=1}^{N} \alpha_k \int_0^1 G_j(r)G_k(r) \, dr = e_j - F_j(m) \qquad j=1,2....N \tag{23}$$

These are just N equations in N unknowns. The G's are linearly independent (if not, the data are l.d., hence one value can be written as a combination of the remainder and provides no new information, reduce N until the G's are l.i.). Thus the matrix formed from values on the l.h.s of eq(23) is non- singular and there is a unique solution for α_j and hence for δm by eq(21).

If the problem is linear and hence eq(15) is exact, the perturbation δm will result in a new model m + δm which satisfies the data (and is closest in the above sense to m). If the problem is non-linear, so that (15) is an approximation to (11) the new model will hopefully lead to a better fit with the data (though this is not guaranteed). If the model m + δm is better it may be used as the starting model in a further step and in this way an iterative solution could be developed providing a systematic method for generating a model whose responses satisfy the data e_j .

2.2 Model interpretation

We now have a model m which fits the data, i.e.

$$F_j(m) = e_j \qquad j=1,2,....N \tag{24}$$

that is values of the functionals F_j calculated from the model m

agree with the measured data e_j. There are an infinity of other models that also fit the data, in particular

$$F_j(m_E) = e_j \qquad j=1,2,\ldots N \tag{25}$$

where m_E is the real earth. What can we deduce about the real earth m_E from our model m?

The model obtained is simply one distibution of the parameter m(r) in (0,1). At a particular radius r_0, the model has a certain value, $m(r_0)$. Is this true of the real earth? - not necessarily, otherwise (24) has a unique solution, so we know there are many other models with different values of $m(r_0)$. What can we say? Suppose we take an average value (maybe a weighted average) of m around r_0. Suppose in particular we take a length ℓ and average m over that length. We might ask the question, what length ℓ (at r_0) do we have to take to ensure that the average over ℓ is the same for <u>all</u> models that satisfy the data (including the real earth)? If we can find such an ℓ then the average of our model over ℓ is some definite property of the earth. If we make ℓ the smallest length over which an average of any model is the same, then we can think of ℓ as the resolving length. Any detail that is finer than ℓ in any given model will be smoothed out by the average and will not be a property of all the models - hence such detail is not resolvable by the data. Our problem is to find ℓ at each radius r_0. If the smallest length scale we can find is very large, then we haven't defined the model parameter very well, we have poor resolution. On the other hand, if we only have to average over a small interval to bring all the models into line, our resolution is good. The deductions we are able to make depend on whether the functionals are linear or non-linear.

2.2.1 Linear functionals

Since linear averaging is simplest, consider a linear average of m to give an estimate $\langle m\rangle(r_0)$, then

$$\langle m\rangle(r_0) = \int_0^1 A(r,r_0)\, m(r)\, dr \tag{26}$$

where A is a weighting function, or averaging kernel. The total weighting must be unity, hence

$$\int_0^1 A(r,r_0)\, dr = 1 \tag{27}$$

Now $F_j(m)$ is l.d. on m(r) and $\langle m\rangle(r_0)$ is l.d. on m(r). It therefore follows that $\langle m\rangle(r_0)$ is l.d. on $F_j(m)$ so that

$$\langle m\rangle(r_0) = \sum_{j=1}^{N} a_j(r_0)\, F_j(m) \tag{28}$$

for some coefficients $a_j(r_0)$, j=1,2,...,N. But since m satisfies the data

$$\langle m\rangle(r_0) = \sum_{j=1}^{N} a_j(r_0)\, e_j \tag{29}$$

so that the average value depends directly on the data. Since the average is independent of the model and comes directly from the data, it is the same for all models satisfying the data. How do we find the appropriate coefficients $a_j(r_0)$? This is the part played by A, the averaging kernel. We want to average m(r) over a small length ℓ around r_0, not over the whole range (0,1). The a_j's must therefore be chosen so that $A(r,r_0)$ is concentrated around r_0, in other words has large values near r_0 and small values away from r_0. Ideally we would like A to be the Dirac delta function $\delta(r-r_0)$, for then $\langle m\rangle(r_0) = m(r_0)$, but this is clearly not possible because m is not unique. One possibility is to make A as close as we can to $\delta(r-r_0)$. The width of the region in which we have succeeded in concentrating A is a measure of the resolving length of the data.

From eqns (2), (26) and (29)

$$A(r,r_0) = \sum_{j=1}^{N} a_j(r_0)\, G_j(r) \qquad (30)$$

i.e. A is a linear combination of the data kernels - a different combination for each r_0. Immediately it is apparent that a linear combination of only a few data kernels will probably not produce a good averaging kernel. The situation may improve with more data. To make A concentrated around r_0 we must define something like the peak width of A and minimise it. There are many definitions of the length scale over which a function is concentrated, a convenient one to use here is the spread of A defined by

$$s(r_0,A) = 12 \int_0^1 (r - r_0)^2 A^2(r,r_0)\, dr \qquad (31)$$

Clearly this is only small if A is concentrated near r_0. The problem then is to choose the coefficients $a_i(r_0)$ so as to minimise $s(r_0,A)$ with the one constraint coming from eq(27) as

$$\sum_{j=1}^{N} a_j(r_0) \int_0^1 G_j(r)\, dr = 1 \qquad (32)$$

This minimisation problem must be solved for each r_0.

2.2.2 Solution

Calculate the data kernels $G_i(r)$ and define

$$u_i = \int_0^1 G_i(r)\, dr \qquad (33)$$

$$S_{ij} = 12 \int_0^1 (r - r_0)^2 G_i(r)\, G_j(r)\, dr \qquad (34)$$

Then minimising the spread (31) subject to the constraint (32) is equivalent to minimising

$$s(r_0,A) = \sum_{i,j=1}^{N} a_i S_{ij} a_j = \underline{a}^T \underline{S}\, \underline{a} \qquad (35)$$

subject to the constraint

$$\sum_{j=1}^{N} a_j u_j = 1 = \underline{u}^T \underline{a} \qquad (36)$$

Introducing a Lagrange multiplier 2λ the minimisation gives

$$\underline{S}\, \underline{a} = \lambda\, \underline{u} \qquad (37)$$

The solution to these N linear equations is

$$\underline{a} = \lambda\, \underline{S}^{-1}\, \underline{u} \qquad (38)$$

where

$$\lambda = 1 / \underline{u}^T \underline{S}^{-1} \underline{u} \qquad (39)$$

This is the algebraic problem which must be solved numerically. With $\underline{a}$ determined for each r_0, s and $\langle m \rangle$ can be calculated at each r_0.

2.2.3 Geometrical interpretation

Backus and Gilbert(1970) give a simple interpretation of the above by considering the case N=3. The G_i's are l.i., so $\underline{S}$ is positive definite. For a fixed s, (35) represents an ellipsoid, (36) a plane. If they intersect, the 2-D figure of intersection in the plane is an ellipse. Take s large enough so that they do intersect, then shrink s. The ellipse of intersection reduces to a single point. With s any smaller, there is no solution. Hence this value of s is the smallest possible spread (s_{min}) and is the solution to the problem, together with $\underline{a}_s$ the point at which it takes this value. Note that a vector normal to the plane is $\underline{u}$ and a vector normal to the surface of the ellipsoid is $\underline{S}\, \underline{a}$. When

$s = s_{min}$ these two are parallel as expressed by (37)

2.3 Errors in data

Up to this stage the data e_j have been assumed perfectly accurate whereas in reality this is never so, and we could not expect there to be any model which gives exact agreement in the form of (17). If σ_j is the one standard error estimate for the datum e_j, for model generation the constraint (17) becomes the inequality constraint

$$e_j - \sigma_j \leq F_j(m+\delta m) \leq e_j + \sigma_j \qquad (40)$$

The model generation calculations proceed as above until this constraint is satisfied.

For model interpretation, suppose the actual error in e_j to be Δe_j. Then the error in the estimate of the average model from (29) is

$$\Delta\langle m\rangle(r_0) = \sum_{j=1}^{M} a_j(r_0)\,\Delta e_j \qquad (41)$$

Choosing tha a_j's so as to minimise the spread would lead to an unacceptably large error as given by (41). It might be better to accept a larger spread if this reduced our error estimate. Naturally the errors themselves, Δe_j, are unknown but we may know something of their underlying statistics. The e_j's may have been obtained as the means of repeated measurements in which case the Δe_j's have zero means $\overline{\Delta e_j} = 0$. Repeated measurements also enable the variance matrix $\underline{E}$ to be determined, with elements $E_{ij} = \overline{\Delta e_i \Delta e_j}$. It will often be that $\underline{E}$ is diagonal with elements σ_j. From (40) an estimate of the variance in the model average $\langle m\rangle(r_0)$ is

$$\varepsilon^2 = \left(\Delta\langle m\rangle(r_0)\right)^2 = \sum_{i,j=1}^{M} a_i E_{ij} a_j \qquad (42)$$

or

$$\varepsilon^2(\underline{a}) = \underline{a}^T \underline{E}\, \underline{a} \qquad (43)$$

so that ε is an estimate of the error in the average value

We have previously determined that $\underline{a}$ which gives the minimum spread, that is the best resolution. That $\underline{a}$ will give an error which is unacceptably large. We could find an $\underline{a}$ which minimises ε^2 - then the spread would be unacceptably large. Backus & Gilbert(1970) prove that decreasing the spread increases the error - a situation analagous to the uncertainty principle.

The problem can be formulated as follows. Consider all vectors $\underline{a}$ for which the spread is a certain value, so that

$$\underline{a}^T \underline{S} \underline{a} = s \qquad (44)$$

Then we ask what $\underline{a}$ that satisfies (44) gives the minimum value of ε^2? The problem is then to minimise (43) with the constraints (44) and (36). For every s we obtain the minimum possible error $\varepsilon(s)$. A graph of spread against error is called a trade-off curve.

A parameterisation sets the problem succinctly. Consider minimising the linear combination of s and ε^2 given by

$$q = \theta\, s + (1-\theta)\, \varepsilon^2 \qquad (45)$$

where the parameter θ varies between 1 and 0. $\theta = 1$ corresponds to minimising spread. $\theta = 0$ corresponds to minimising error. As θ goes from 1 to 0 the trade-off curve $\varepsilon(s)$ is described and this, as Backus & Gilbert prove, is a monotonically decreasing function of s.

Again examine the case for N=3. Eq(44) represents an ellipsoid and since $\epsilon^2 > 0$, (43) represents an ellipsoid. These two intersect the plane (36) as ellipses. If we decrease s the region of intersection shrinks to a point where $\underline{a} = \underline{a}_s$ and this gives the minimum possible spread s_{min} - here $\theta = 1$. Decreasing ϵ^2 we find a point $\underline{a}_\epsilon$ which gives the minimum possible variance ϵ^2_{min} - at this point $\epsilon = 0$. For a fixed s (or θ) there will be a point where the ellipses just touch. This point, say $\underline{a}(s)$ is the solution to the problem - it gives the minimum ϵ^2 for a given s. As s increases from s_{min}, $\underline{a}(s)$ traces a path from $\underline{a}_s$ to $\underline{a}_\epsilon$. There is no point in increasing s further as there will be no further reduction in error.

In the earth some physical properties, for example conductivity, vary over many orders of magnitude - it might then be more appropriate to consider relative errors in the model rather than absolute errors. The geometry changes to that of the intersection of a plane with a double cone. This can result in an ellipse, a parabola or a pair of hyperbolae and each case has to be examined separately in conjunction with the intersection of the spread ellipsoid with the plane. It is sometimes easier in these cases to define the logarithm of the parameter as the model and then consider the somewhat easier case of absolute errors.

In the above we have assumed linear functionals. This assures us that the calculated averages and inferences we make from our single determined model are also true of the real earth.

If the functionals F_i are non-linear the computations can proceed almost without change for their linearised approximation - however the inferences we can make are no longer precise - we can no longer make definite statements concerning properties of the earth. This is indeed a high price to pay.

For non-linear functionals eqns (28) and (29) are no longer true. However it is still of use to construct linear averages of the model m of the form

$$\langle m \rangle(r_0) = \sum_{i=1}^{N} a_i(r_0)\, q_i \qquad (46)$$

where

$$q_i = (G_i, m) \qquad (47)$$

The difference in calculation is the use of q_i in place of the data e_i. Now the model average will depend on the model m - the trade-off curve will also be model dependent. Perhaps we could have chosen a different starting model that would have led to a model m' satisfying (24). This model could be quite different from m, with different average properties. We cannot be certain of earth properties until all the models satisfying (24) are determined - but that is beyond us for there are an infinity of them. It is often possible to use other data, or pre-conceived ideas, to limit the form of our models but the limitations inherent in linearised rather than linear problems needs always to be remembered.

2.5 Computational details

2.5.1 Model finding

The matrix formed from (22) or (23) is theoretically non-singular but in practise the equations may be ill-conditioned, especially when the matrix is of large dimension. An alternative approach to direct matrix inversion, developed by Parker(1977), is to consider a spectral expansion of the model m. Eqn (22) may be written

$$\underline{\Gamma}\ \underline{\alpha} = \underline{\gamma} \tag{48}$$

where the elements of $\underline{\Gamma}$ are given by

$$\Gamma_{ij} = \int_0^1 G_i\, G_j\, dr \tag{49}$$

the elements of $\underline{\alpha}$ are α_i and the elements of $\underline{\gamma}$ are $e_i - F_i(m)$. Since $\underline{\Gamma}$ is positive definite and symmetric it may be diagonalised with an orthogonal matrix $\underline{Q}$ such that

$$\underline{Q}^T\ \underline{\Gamma}\ \underline{Q} = \underline{\Lambda} \tag{50}$$

where

$$\underline{\Lambda} = \mathrm{diag}(\lambda_1, \lambda_2, \ldots\ldots, \lambda_N) \tag{51}$$

and $\lambda_1 > \lambda_2 > \lambda_3 > \ldots\ldots > \lambda_N > 0$.

Thus eqn(48) transforms to

$$\underline{\Lambda}\ \underline{Q}^T\ \underline{\alpha} = \underline{Q}^T\ \underline{\gamma} \tag{52}$$

with the simple solution

$$\underline{\alpha} = \underline{Q}\ \underline{\Lambda}^{-1}\ \underline{Q}^T\ \underline{\gamma} \tag{53}$$

The orthogonalisation may be achieved by singular value decomposition. With this orthogonalisation eqn(21) becomes

$$\delta m = \sum_{j=1}^{N} \alpha_j \left(\lambda_j^{-1/2} \sum_{i=1}^{N} Q_{ij}\, G_i \right) \tag{54}$$

It will often be the case that many of the eigenvalues λ_j are small and closely spaced leading to instability. With the eigenvalues arranged in decreasing order it is preferable now to truncate eqn(52) leaving the first few, say L, large eigenvalues only. The orthogonalisation actually has the effect of combining the data into "significant" pieces of information (associated with the larger eigenvalues). The truncation effectively takes only terms j = 1 to L in the spectral expansion (54). Parker(1977) shows further that a suitable truncation level L may be determined by examining the misfit χ^2 between the data e_i and substitution of the truncated form of (54) into $F_j(m+\delta m)$. For L terms retained, the misfit is

$$\chi^2 = \sum_{j=L+1}^{N} \lambda_j\, e_j^2 \tag{55}$$

Coefficients associated with small eigenvalues contribute little to the misfit but greatly to the solution and its uncertainty, the standard error in each α_j being $\lambda_j^{-1/2}$.

2.5.2 Model interpretation

Trade-off calculations either for linear or non-linear problems and including data errors require solution of

$$\underline{W}\ \underline{a} = \lambda\ \underline{u} \tag{56}$$

with

$$\underline{a}^T \cdot \underline{u} = 1 \tag{57}$$

where

$$\underline{W} = \theta \underline{S} + (1-\theta) \underline{E} \qquad (58)$$

$\theta = 1$ corresponds to the error-free algebraic solution detailed earlier.

As above, diagonalisation is the key. Diagonalisation of $\underline{W}$ requires simultaneous diagonalisation of $\underline{S}$ and $\underline{E}$, which can be achieved since both are symmetric and positive definite. $\underline{E}$ is usually diagonal but may not necessarily be. If $\underline{E}$ is diagonal, scale the matrix $\underline{S}$ with $\underline{E}^{-1/2}$ so that

$$\underline{W}' = \underline{E}^{1/2} \underline{W} \underline{E}^{1/2}$$

$$= \theta \underline{E}^{-1/2} \underline{S} \underline{E}^{-1/2} + (1-\theta) \underline{I} \qquad (59)$$

where $\underline{I}$ is the unit matrix and write $\underline{W}'$ as

$$\underline{W}' = \theta \underline{S}' + (1-\theta) \underline{I} \qquad (60)$$

Now diagonalise $\underline{S}'$ (and hence $\underline{W}'$) with an orthogonal transformation $\underline{O}$ so that the elements of $\underline{W}'$ become

$$W_i' = \left(\theta \underline{O} \underline{S}' \underline{O}^T + (1-\theta) \underline{I} \right)_i$$

$$= \theta d_i + (1-\theta) \qquad (61)$$

where d_i are the diagonal elements of $\underline{S}'$. Applying the transformation to $\underline{u}$ and $\underline{a}$ so that $\underline{u}' = \underline{O} \underline{u}$ and $\underline{a}' = \underline{O} \underline{a}$ the system equations become

$$\underline{W}' \underline{a}' = \lambda \underline{u}' \qquad (62)$$

and

$$\underline{a}'^T \cdot \underline{u}' = 1 \qquad (63)$$

Since $\underline{W}'$ is now diagonal, the solution is simply

$$a_i' = \lambda u_i' \left(\theta d_i + (1-\theta) \right)^{-1} \qquad (64)$$

where

$$\lambda = 1 / \sum \left(u_i'^2 \left(\theta d_i + (1-\theta) \right)^{-1} \right) \qquad (65)$$

Calculations for the spread and variance of the estimate are then

$$\text{spread} = \underline{a}^T \underline{S} \underline{a} = \sum a_i'^2 d_i \qquad (66)$$

$$\text{variance} = \underline{a}^T \underline{E} \underline{a} = \sum a_i'^2 \qquad (67)$$

These calculations can be made for each θ and the trade-off between spread and error investigated.

2.6 General procedure for inversion studies

(1) Generate a model that fits the data and associated errors. An acceptable model may be one that satisfies the inequalities (40) or alternatively one that yields a suitably small value of misfit χ^2. The model parameter may not necessarily vary smoothly, there may be fine scale detail superimposed on more general trends.

(2) Calculate and examine the trade-off information and decide on a suitable error. Minimise the spread with this error, resulting in a set of coefficients a_i for each selected radius r_0. Alternatively find the spread and error at the "knee" of each trade-off curve. Compute the model averages $\langle m \rangle$ for each r_0.

(3) Examine the averages together with their error and spread estimates. If, for example, the model and the averages

exhibit fine detail, the associated error and spreads will indicate whether or not such detail is resolvable.

3. Applications

3.1 The inversion of travel time data

Let the earth be spherically symmetric with velocity structure v(r). Then the travel time T for a seismic wave travelling from a surface source to a surface receiver epicentral distance Δ away is given by

$$T(p) = 2 \int_{r_0}^{R} (r/v^2) (r^2/v^2 - p^2)^{-1/2} \, dr \qquad (68)$$

where R is the earth's radius, p is the ray parameter and r_0 is the radius at the bottom of the ray path. The epicentral distance Δ may also be expressed in integral form as

$$\Delta(p) = 2 p \int_{r_0}^{R} r^{-1} (r^2/v^2 - p^2)^{-1/2} \, dr \qquad (69)$$

The ray parameter p is given by

$$p = \frac{dT}{d\Delta} \qquad (70)$$

and satisfies the relation

$$p = v^{-1} r \sin i \qquad (71)$$

along the ray path, where i is the angle between the ray and the radius vector. Thus at the deepest penetration of the ray, where $i = \pi/2$,

$$r_0 = p v \qquad (72)$$

3.1.1 Formal inversion by Herglotz & Wiechert

Travel time data consists of pairs of values (T,Δ) for a particular body wave type. Differentiation enables the (p,Δ) relations to be determined in principle. Eqn (69) specifies $\Delta(p)$ as a function of the velocity structure v(r) - what we require is the inverse, that is v(r) as a function of the (p,Δ) relations. Herglotz(1907) and Wiechert(1910) provide an analytic solution which is presented before the Backus-Gilbert approach is applied.

In eqn(69) put $\eta = r/v$ and change the variable of integration from r to η. Then

$$\Delta(p) = 2 \int_{p}^{R/v_0} p (\eta^2 - p^2)^{-1/2} \left(\frac{1}{r} \frac{dr}{d\eta} \right) d\eta$$

$$= 2 \int_{p}^{R/v_0} p (\eta^2 - p^2)^{-1/2} \frac{d}{d\eta} (\ln r) \, d\eta \qquad (73)$$

where $v_0 = v(r_0)$. This is now an integral equation determining $\ln r$ as a function of η/v.

Suppose η decreases monotonically as r decreases in the range $R \geq r > r'$ for some r' and let $\eta' = r'/v(r')$. Let η_1 be such that $\eta_1 > \eta'$. Then (73) holds for $\eta_1 \leq p \leq R/v_0$. Multiply (73) by $(p^2 - \eta_1^2)^{-1/2} \, dp$ and integrate from $p = \eta_1$ to R/v_0. Then

$$\int_{\eta_1}^{R/v_0} (p^2 - \eta_1^2)^{-1/2} \, \Delta \, dp = 2 \int_{\eta_1}^{R/v_0} dp \int_{p}^{R/v_0} p \left(p^2 - \eta_1^2\right)^{-\frac{1}{2}} \left(\eta^2 - p^2\right)^{-\frac{1}{2}} \frac{d}{d\eta} \left(\ln r\right) d\eta$$

$$= 2 \int_{\eta_1}^{R/v_0} d\eta \int_{\eta_1}^{\eta} p \left(p^2 - \eta_1^2\right)^{-\frac{1}{2}} \left(\eta^2 - p^2\right)^{-\frac{1}{2}} \frac{d}{d\eta} \left(\ln r\right) dp \qquad (74)$$

on changing the order of integration. The integral over p is standard and with $\eta > \eta_1$

$$\int_{\eta_1}^{\eta} p (p^2 - \eta_1^2)^{-1/2} (\eta^2 - p^2)^{-1/2} \, dp = \pi/2 \qquad (75)$$

Therefore

$$\int_{\eta}^{R/v_0} \Delta \, (p^2 - \eta_1^2)^{-1/2} \, dp = \pi \int_{\eta}^{R/v_0} \frac{d}{d\eta} (\ln r) \, d\eta$$

$$= \pi \ln (R/r_1) \qquad (76)$$

where $r_1 = \eta_1 v_1$. The l.h.s. may be integrated by parts yielding

$$\int_0^{\Delta_1} \cosh^{-1} (p/\eta_1) \, d\Delta = \pi \ln (R/r_1) \qquad (77)$$

where $\Delta = \Delta_1$ at $\eta = \eta_1$. This integral equation gives r_1 as a function of η_1 ($= r_1/v_1$) and hence v as a function of r.

Use of this equation is as follows. p is a known function of Δ (the p,Δ relations). Select an epicentral distance Δ_1 and determine $\eta_1 = dT/d\Delta$ at Δ_1 (i.e. the slope of the travel time curve at Δ_1). Now the l.h.s. of (77) can be evaluated by determining p at intermediate points and evaluating $\cosh^{-1}$ (p/η). Hence r_1 is determined. But $r_1 = \eta_1 v_1$ thus determining v_1. This procedure holds for any r_1 in the range $R \geq r_1 > r'$. Hence v is determined as a function of r down to radius r' (i.e. provided η is monotonically decreasing).

Complications arise when η increases. For a discontinuous increase in η (implying a discontinuous decrease in velocity as at the mantle/core boundary) shadow zones are created and there are gaps in the (T,Δ) relations. Triplications occur in the (T,Δ) relations for a rapid increase in velocity with depth.

3.1.2 The function Tau - p

The above classic methods of Herglotz & Wiechert have been extended by Gerver & Markusevitch (1966) to allow for the presence of a finite number of low velocity zones. All these classical methods have disadvantages. First they require knowledge of the complete exact travel time curve, requiring in practice interpolation. Second, complexities in the (T,Δ) or (p,Δ) relations make these functions difficult to estimate from

measurements. The functions may not be single valued (as at triplications) and may only be partially defined (as in the case of shadow zones).

The problem may be circumvented by considering the single valued delay time $\tau(p)$ defined as

$$\tau(p) = T(\Delta) - p \, \Delta(p) \qquad (78)$$

From (68) and (69)

$$\tau(p) = 2 \int_{r_0}^{R} r^{-1} (\eta^2 - p^2)^{1/2} \, dr \qquad (79)$$

τ has the useful property of being monotonic. Differentiating (78)

$$\frac{d\tau}{dp} = -\Delta \qquad (80)$$

hence, since Δ is always positive, τ is a monotonic decreasing single valued function of p. A triplication is therefore "unwrapped" and appears on a $\tau(p)$ plot as a kink. Low velocity regions result in discontinuities in τ but no discontinuity in p.

Discrete values of the function τ_j at points p_j for j=1,2,.....,N together with estimates of error bounds may be determined by the methods of Bessonova et al.(1974) or Kennett(1976). These form the discrete data set to which a (slightly modified) Backus-Gilbert inverse scheme may be applied.

As with many examples in geophysics, there is an equivalence between spherical and planar geometry. The above spherical earth problem may be converted to its planar equivalent in $z > 0$ by the transformation

$$x = R \, \Delta$$

$$z = R \ln (R/r) \qquad (81)$$

$$u(z) = (r/R) \, (1/v(r))$$

x is now distance along the surface, z is depth and u(z) is the

slowness profile (reciprocal of the velocity). With this transformation eqn(79) becomes

$$\tau(p) = 2 \int_0^{z_p} \left(u^2(z) - p^2 \right)^{1/2} \, dz \qquad (82)$$

where z_p is the maximum depth reached by the ray with parameter p.

3.1.3 The inversion of $\tau(p)$

It is clear that $\tau(p)$ depends non-linearly on the slowness parameter $u(z)$ so that the linearised version of (82) is required. A small perturbation δu will lead to a small change $\delta\tau$ according to

$$\delta\tau(p) = 2 \int_0^{z_p} u(z)\, \delta u(z) \left(u^2(z) - p^2 \right)^{-1/2} \, dz \qquad (83)$$

$$= 2 \int_0^{z_p} u^2 \left(u^2 - p^2 \right)^{-1/2} m(z) \, dz \qquad (84)$$

where the model m is introduced as the relative perturbation

$$m(z) = \delta u(z)/u(z) \qquad (85)$$

Now let z_M be the greatest depth to be considered in any model. Then (84) may be written

$$\delta\tau_i = \int_0^{z_M} G_i(z)\, m(z) \, dz \qquad (86)$$

where

$$G_i(z) = 2 u^2 \left(u^2 - p_i^2 \right)^{-1/2} \qquad 0 \leq z < z_p$$

$$= 0 \qquad z \geq z_p \qquad (87)$$

Having now identified the Fréchet derivative we would like to be able to proceed to the calculations detailed above. Model

interpretation does proceed along these lines but the preceding step of model generation has one further complexity. In eqn(49) the Fréchet kernel is required to be square integrable, but from (87) this is not so. The singularity in G may be removed by integrating by parts:

$$\int_0^{z_M} G_i(z)\, m(z) \, dz = J_i(0)\, m(0) + \int_0^{z_M} J_i(z)\, m'(z) \, dz \qquad (88)$$

where

$$J_i(z) = \int_z^{z_p} G_i(s) \, ds \qquad (89)$$

and m' is the derivative of m. The new kernels J are the travel times from level z to the bottoming of the ray z_p and thus may be calculated for any given velocity structure. We now seek the smoothest perturbation, that is the minimum of

$$\int_0^{z_M} \left(m'(z) \right)^2 dz + m^2(0) \qquad (90)$$

subject to the constraints (40). Having found m(0) and m' the latter may be integrated to yield the relative perturbation m(z) and a "new" model determined. Since this is a non-linear problem, the new model could be used as the starting model in the next step, thus generating an iterative process.

3.2 The inversion of free oscillation data

The problem of determining an earth model (density and seismic velocities) satisfying measured squared eigenfrequencies of the earth's spheroidal and toroidal free oscillations, in part stimulated the inversion theory developed by Backus and Gilbert. Before their theory, although forward calculations could be performed, it was not clear how to change a model to give better agreement with the substantial body of data available.

An earth model consists of the 3-tuple (ρ,κ,μ) where ρ is density and κ and μ are the bulk and shear moduli respectively. Thus for the application of inverse theory three Fréchet derivatives are required corresponding to changes in ρ, κ and μ. Rayleigh's principle is first invoked to yield the squared eigenfrequency ω^2 for a particular normal mode. Applying small perturbations $\delta\rho$, $\delta\kappa$ and $\delta\mu$ and linearising determines the required kernels K,M,R which satisfy

$$(\delta\omega^2) \int_V \rho\, s^2\, dv = \int_V (K\delta\kappa + M\delta\mu + R\delta\rho)\, dv \qquad (91)$$

where V is the volume of the earth and $\underline{s}$ the displacement field. The form of K, M and R, together with their derivations are given explicitly in Backus and Gilbert(1967).

The squared frequencies of the normal modes, which are non-linear functionals, are augmented by the linear functionals of mass and moment of inertia to yield a data set for inversion. The simplest application is to assume the compressional and shear wave velocities (and hence κ and μ) to be known and to infer the density ρ as a function of radius r. Early experiments on small data sets illustrated the nature of the problem and solution and were given by Backus and Gilbert (1967, 1968, 1970). Dziewonski and Gilbert(1972) extended the data set with the inclusion of 70 overtones (spheroidal and toroidal) together with fundamental modes. Some of the overtones are interpreted as implying solidity of the inner core. Gilbert and Dziewonski(1975) have given a standardised data set consisting of 1064 distinct eigenfrequencies and have applied these to refine earth models and to determine the mechanism of an earthquake source.

3.3 The inversion of geomagnetic data

Geomagnetic data may be inverted to yield the earth's conductivity profile $\sigma(r)$. Taking the earth as spherical, outside any conducting region the magnetic scalar potential Ω satisfies Laplace's equation

$$\nabla^2\Omega = 0 \qquad (92)$$

In conducting regions the electric field $\underline{E}$ satisfies the diffusion equation

$$\nabla^2\underline{E} = \mu\sigma \frac{\partial \underline{E}}{\partial t} \qquad (93)$$

where μ is the permeability of free space. The appropriate solution of (92) may be written

$$\Omega = \sum_{n=0}^{\infty} \sum_{m=0}^{n} \{ a_n^m r^n + b_n^m r^{-n-1} \} S_n^m(\theta,\phi) \qquad (94)$$

where S_n^m is a surface harmonic of degree n order m and a and b represent external (inducing) and internal (induced) parts of the potential respectively. In the conducting regions, the appropriate solution of (93) is

$$\underline{E} = \text{curl}\, \psi \underline{r} \qquad (95)$$

where

$$\psi = \sum_{n=0}^{\infty} \sum_{m=0}^{n} R_n^m(r)\, S_n^m(\theta,\phi) \qquad (96)$$

with R_n^m satisfying a second order differential equation

$$\frac{d}{dr}\left(r^2 \frac{d}{dr} R_n^m\right) + \{ i\omega\mu\sigma r^2 - n(n+1) \} R_n^m = 0 \qquad (97)$$

A number of 'response measures' may be constructed, for example

$$Q_n^m = b_n^m / a_n^m \qquad (98)$$

which is the ratio of internal to external parts. Another such measure is the logarithm of apparent resistivity given by

$$\gamma = \ln (i\omega\mu c^2) \qquad (99)$$

where c, the penetration depth, is

$$c = a (1 - 2 Q) / 2 (1 + Q) \qquad (100)$$

and the sub- and super-scripts have been dropped from Q.For the response γ the Fréchet derivative may be determined as

$$F(\gamma,r,\omega) = - 2 r^2 \mu i\omega R^2(r) / R(1) R'(1) \qquad (101)$$

Clearly the problem is non-linear since F depends on σ through the function R. Parker(1970) applies this theory to determine the electrical conductivity in the earth's mantle and Hobbs(1977) and Hobbs et al.(1984) give applications concerning the moon.

4 Conclusions

Provided the forward problem can be solved and the appropriate Fréchet derivative found, inverse theory provides a systematic method for determining an earth model satisfying data and for interpreting that model in terms of its linear averages. The severe limitations inherent in the use of linearised versions of non-linear problems must always be remembered

References

Backus,G. & Gilbert,F.,1967. Numerical applications of a formalism for geophysical inverse problems, Geophys.J.R.astr.Soc.,13,247-276.

Backus,G. & Gilbert,F.,1968. The resolving power of gross earth data, Geophys.J.R.astr.Soc.,16,169-205.

Backus,G. & Gilbert,F.,1970. Uniqueness in the inversion of inaccurate gross earth data,Phil.Trans.R.Soc.Lond.,A266,123-192.

Bessonova,E.N.,Fishman,V.M.,Ryaboyi,V.Z. & Sitnikova,G.A.,1974. The tau method for the inversion of travel times - I Deep seismic sounding data, Geophys.J.R.astr.Soc.,36,377-398.

Dziewonski,A.M. & Gilbert,F.,1972. Observations of normal modes from 84 recordings of the Alaskan earthquake of 1964 March 28, Geophys.J.R.astr. Soc.,27,393-466.

Gerver,M. & Markusevitch,V.,1966. Determination of seismic wave velocity from the travel time curve, Geophys.J.R.astr.Soc.,11,165-173.

Gilbert,F. & Dziewonski,A.M.,1975. An application of normal mode theory to the retrieval of structural parameters and source mechanisms from seismic spectra, Phil.Trans.R.Soc.Lond.,A278,187-269.

Herglotz,G.,1907. Uber das Bunndorfsche Problem der Fortpflanzungsgeschwindigkeit der Erdbebenstahlen, Phys.Z.,8,145-147.

Hobbs,B.A.,1977. The electrical conductivity of the moon: an application of inverse theory, Geophys.J.R.astr.Soc.,51,727-744.

Hobbs,B.A.,Hood,L.L.,Herbert,F. & Sonett,C.P.,1984. Low frequency electromagnetic induction in the moon: linearised inverse theory and lunar core calculations, Geophys.J.R.astr.Soc.,79,691-696.

Kennett,B.N.L.,1976. A comparison of travel-time inversions, Geophys. J.R.astr.Soc.,44,517-536.

Parker,R.L.,1970. The inverse problem of electrical conductivity in the mantle, Geophys.J.R.astr.Soc.,22,121-138.

Parker,R.L.,1977. Understanding inverse theory, Ann.Rev.Earth Planet. Sci.,5,35-64.

Weichert,E.,1910. Bestimmung der weges der Erdbebenwellen im

Erdinnern, I, Theoretisches, Phys.Z.,11,294-304

Suggested further reading

R.Cassinis(Editor), "The solution of the inverse problem in geophysical interpretation", Plenum Press, 1981.

W.Menke, "Geophysical data analysis: Discrete inverse theory", Academic Press, 1984.

S.Twomey, "Developments in geomathematics 3: Introduction to the mathematics of remote sensing and indirect measurements", Elsevier, 1977.

B.A.Bolt(Editor), "Methods in computational physics Vol 12 Seismology: Body waves and sources", Academic Press, 1972. (See chapter by Johnson and Gilbert on Inversion and inference for teleseismic ray data).