[29] W.D.Dozier, J.M.Carpenter and G.P. Felcher, A New Method for Phase Determination in Neutron Reflectivity Measurements, Bulletin of APS, Vol.36, No.3, (1991), 772.

[30] L.D. Landau and E.M. Lifshits, Quantum Mechanics, 3rd ed., Pergamon Press, (1977), 164-189

INVERSE PROBLEMS IN SEISMOLOGY

J TRAMPERT, J-J LEVEQUE AND M CARA

ABSTRACT

Seismologists have been confronted with inverse problems for a long time. At the beginning of the century, Herglotz and Wiechert showed that, under certain assumptions, the seismic velocity distribution with depth could be uniquely determined from travel time data. With this method, Jeffreys proposed in 1939 the first practical and still widely used Earth model. The seismological inverse problem took a major turning with the works of Backus and Gilbert at the end of the sixties. They introduced the concepts of resolution and trade-off between error and resolution, setting thus the basis for a more theoretical approach to the inverse problem in seismology. Franklin (1970) introduced in a stochastic approach the covariance operators for gaussian probability densities, and Jackson (1979) extended these results to non-gaussian densities. Wiggins (1972) first applied the singular value decomposition to a seismological inverse problem. Finally, a more general statement of the inverse problem has been proposed by Tarantola and Valette in 1982. With the effort to describe the Earth as precisely as possible, new problems arose, namely the trade-off between different model parameters and the inconsistency problem. Seismic tomography is after the hypocenter determination the most common inverse problem in seismology. Waveform inversion, while containing the most possible information, needs generally a carefully chosen formulation in order to deal with the more complicated relationship between data and model. The introduction of an inexact theory relating data and model can be very interesting for non-deterministic formulations. We show recent examples to illustrate these different techniques and difficulties going with them and, as a conclusion, point out current research topics of inverse problems in seismology.

1. INTRODUCTION

At the beginning of the century, seismologists were first confronted with inverse problems. As soon as it was understood that recordings could be related to earthquakes happening thousands of kilometers away (fig. 1), they tried to use observed travel time curves (travel time as a function of distance) to retrieve the seismic velocity distribution of the Earth's interior. This non-linear inverse problem was simultaneously solved by Herglotz (1907) and Wiechert (1907). Under the assumption of a monotonous increase of the relevant parameter (Earthradius/velocity) with depth, the method yields a unique and analytical solution of the velocity as a function of depth.

Jeffreys (1939) applied this technique to all available data at that time to compute an Earth model which is still widely used as a reference model (fig. 2).

In the last 20 years, inverse problems have been intensively studied in seismology. It has become such a vast subject that a complete review of its importance and application in seismology would need a book on its own. Therefore, we have chosen to focus on some illustrative problems of inverse theory encountered in seismology.

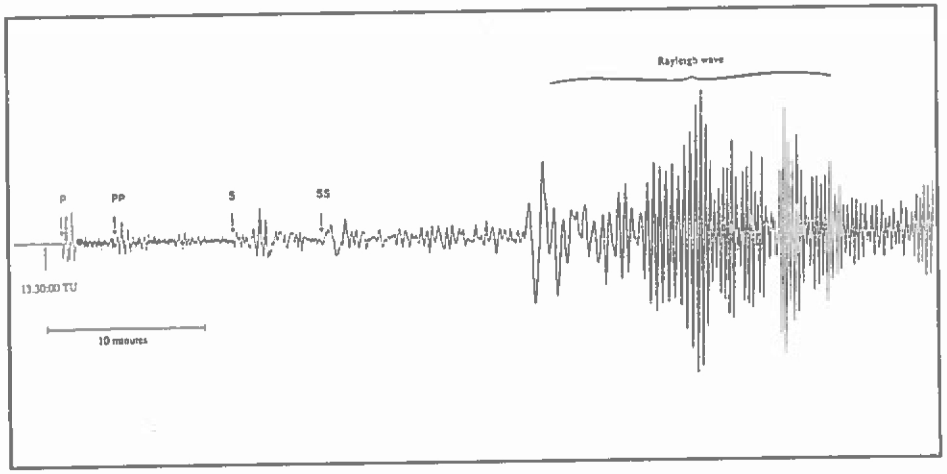


Figure 1: Seismogram, recorded in Strasbourg, of the 19 September 1985 Michoacan earthquake, which caused extensive damage in Mexico City.

The seismogram shown in figure 1 has some characteristic features (such as P and S body waves and dispersed surface waves) which are the result of the seismic source, the constitution of the Earth through which the waves propagate and the characteristics of the seismic recorder. Due to the linearity of the problem of elastic wave propagation, the Earth's displacement u(t) seen at a seismic station can be written as

$$u(t) = s(t) * p(e,t) * i(t) \qquad (1)$$

where s(t) describes the source mecanism, p(e,t) the propagation in the given Earth model e and i(t) the instrument response. The unknowns are typically the source s, and the Earth structure e. The aim of seismology is to extract data from such seismograms which allow to solve for the chosen unknowns. Formally

$$\begin{pmatrix} s \\ e \end{pmatrix} = h(u,p,i) \qquad (2)$$

where h is the inverse operator. Ideally, we should invert the whole seismogram to obtain s and e, but this still presents considerable theoretical and technical problems.

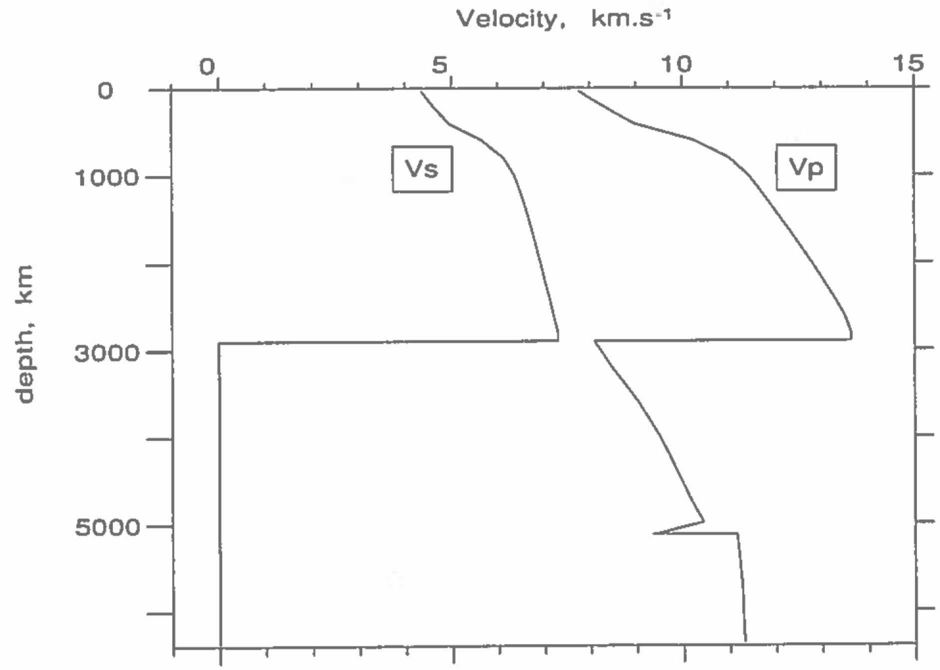


Figure 2: Earth model after Jeffreys, 1939.

Seismologists prefer to partition the problem in two different ways. First, in seismic source studies, the Earth structure is generally considered to be perfectly well known, while the seismic source is fixed for Earth structure inversions. Second, the seismic records (data) are commonly partitionned as well. First arrival times of P and S waves are used to compute the daily hypocenter determinations of the recorded earthquakes. In seismic tomography, these arrival times compared to theoretical arrival times give 3D velocity anomalies around a reference model (Dziewonski, 1984). The waveforms of body waves are used to infer focal mecanisms (Kanamori and Cipar, 1974; Langston and Helmberger, 1975; Deschamps et al., 1980) and the detail of the seismic rupture processes (Das and Kostrov, 1990). Other interesting observables are dispersion curves (velocity as a function of frequency) of surface waves. They give good information on the shallow Earth structure such as crust and upper mantle (Wiggins, 1972; Montagner and Nataf, 1988). The waveforms of surface waves are mainly used for the study of the mantle structure (Woodhouse and Dziewonski, 1984;

Tanimoto, 1987; Nolet, 1987; Cara and Leveque, 1987). Free oscillations, a standing wave pattern, excited by strong earthquakes, put the whole Earth into vibration and give long wavelength heterogeneities on a global scale (Woodhouse et al., 1986).

Most of the inverse problems treated in seismology are linear or linearized and minimize an L_2 norm. With this assumption, we present the most important concepts, show some recent applications and discuss current research activity.

2. GENERAL CONCEPTS OF LINEAR INVERSE PROBLEMS

In matrix form, the basic equation for a linear discrete problem is

$$d = Gm \tag{3}$$

where **d** are the measured data and **G** describes the theory operating on the model parameters **m**. The inverse problem consists in finding **m** from **d**. In the case where G is a non-singular square matrix,the solution of (3) is obvious. The general case is not as easy to solve. G may be either square and singular, or non-square, or errors in the data make that relation (3) is not exact. The solution may then be given by the classical least squares where several cases need to be distinguished:

(i) We have enough information to evaluate all model parameters, but some equations are contradictory due to measurement errors. The problem is then said to be purely overdetermined and G^tG is regular, which yields the solution

$$m^* = (G^tG)^{-1}G^td \tag{4}$$

where the L_2 norm $\|d-Gm\|$ has been minimized.

(ii) There are no contradictions in the available informations, but we don't have enough equations to evaluate all the model parameters. The problem is said to be purely underdetermined and GG^t is regular. The solution is

$$m^* = G^t(GG^t)^{-1}d \tag{5}$$

and the model fits exactly the data ($\|d-Gm\| = 0$). It should be noted that in this case, an infinite number of models exactly fit the data, and the model chosen by (5) is the one with a minimum norm. With erroneous data, this last solution is physically unacceptable. As the predicted model exactly fits wrong data, m^* is necessarily wrong.

(iii) Finally, the last and most common case, we have contradictory information on some model parameters, and at the same time some other parameters cannot be assessed due to a lack of information. In order to smooth the effect of data errors on the predicted model, Levenberg (1944) introduced a damped least squares solution given by

$$m^* = (G^tG+\theta^2I)^{-1}G^td$$

$$= G^t(GG^t+\theta^2I)^{-1}d \tag{6}$$

where θ^2 is a constant. This algorithm can be applied regardless the nature of G.

Franklin (1970) formalized the intuitive notion of the damping parameter in terms of covariance operators over the model parameters and the data. His stochastic view assumes gaussian probability densities. Jackson (1979) extended Franklin's results to the case of non-gaussian probability densities which includes the case of rigid bounds. A more general approach has been described by Tarantola and Valette (1982 a,b) and can be applied with any probability density function as well as to non-linear problems.

2.1. The Backus-Gilbert method

Since Jeffreys' Earth model (1939) until the early sixties, seismologists have hardly done any other inverse problems than locating earthquakes using equation (4). In a series of papers, Backus and Gilbert (1967, 1968, 1970) drew the seismologists' attention to a more general way of addressing the inverse problem. To solve (3), they want the estimated model m^* to be a linear combination of the data:

$$m^* = Hd \tag{7}$$

using (3), they may write that

$$m^* = HGm = Rm \tag{8}$$

R is called the resolving kernel and describes a filter through which the true model **m** is seen. We can arbitrarily choose the operator **H**. In the case of perfect data, Backus and Gilbert introduced the J-deltaness criterion, for example, which consists in choosing **H** in a way that **R** is closest to the identity operator and they have to minimize

$$\|R-I\| \tag{9}$$

In the case of noisy data, they choose the coefficients of H in such a way that the estimated model is optimal with respect to the resolution kernel and the estimated model error. In this case, the minimization involves the quantity

$$(1-\alpha)\|R-I\| + \alpha \,\mathrm{tr}\,(C_{m^*}) \tag{10}$$

where $0 \le \alpha \le 1$ determines the trade-off between error and resolution. The better the resolution of the final model, the higher its error and vice versa. The estimated model still exactly fits the data and we have already seen that this is physically unacceptable. Gilbert (1971) extended the method to the case where the data are to be fitted only within the error bars, which is physically more satisfying.

2.2. The generalized least squares

Tarantola and Valette (1982 a,b) following a stochastic approach, formulated the inverse problem in very general terms using different states of information described by probability densities. Using Bayes' theorem, they combine the different states of information to obtain a unique solution. If the probability laws describing these states of information are gaussian, the generalized least squares formalism is obtained. Assuming that there are no correlations between the data and the model parameters, and that the theory G is perfectly well known, their formalism reduces to the stochastic inverse described by Franklin (1970), where the minimized cost function is

$$\|d - Gm\|_{C_d^{-1}} + \|m - m_0\|_{C_m^{-1}} \tag{11}$$

which gives the solution

$$m^* = m_0 + (G^tC_d^{-1}G + C_m^{-1})^{-1}G^tC_d^{-1}(d-Gm_0)$$

$$= m_0 + C_mG^t(GC_mG^t + C_d)^{-1}(d-Gm_0) \tag{12}$$

with m_0 the a priori model, C_m the covariance operator describing the error ellipsoid around m_0 and C_d the covariance operator corresponding to the data. It is easily seen that (6) is a special case of (12) (e.g. Aki and Richards, 1980). Expressions (4) and (5) are not special cases of (12), but only limit cases of (12). Indeed, due to the definition of covariance operators, neither C_d nor C_m^{-1} can be $\mathbf{0}$. Expression (11) reflects a trade-off between the data misfit and the model variation. For a given metric C_d^{-1}, the metric C_m^{-1} can be thought of a parameter describing the trade-off. Imagine the case where you have a good a priori model and you are not too sure about the quality of your data. A high C_m^{-1} will force the estimated model to stay close to m_0. For good quality data, the high C_d^{-1}, will give the most weight on the data misfit.

On the other hand it can be shown that minimizing (11) is equivalent to minimizing the trace of the covariance matrix C_{m*} of the estimated model (Tarantola, 1987). As

$$C_{m*} = (I-R)C_m \tag{13}$$

modifying C_m describes then a trade-off between error and resolution similar to the one introduced by Backus and Gilbert.

2.3. Other methods

A different approach to the linear inverse problem, based on the singular value decomposition, has been developped by Penrose (1955). He shows that each matrix G can be associated to a generalized inverse, which is the classical inverse G^{-1}, if G is regular. Lanczos (1961) describes how to build this inverse and shows that this method consists in projecting the real data to the sub-space of all possible data G{m}, and to choose among all possible models which exactly fit the data, the model which has no component in the sub-space associated with the zero eigenvalues. The solution is equivalent to (5). Wiggins (1972) proposed a modification of Lanczos' method to obtain a physically acceptable solution in the case of noisy data. The stabilization of the final model is obtained by truncating the eigenvector bases. Wiggins' inverse is similar to the damped least squares (6). Finally , the generalized inverse obtained by spectral decomposition as described by Matsu'ura and Hirata (1982) establishes the relationship with the stochastic inverse of Franklin (1970).

2.4. Specific problems due to multi-parameter inversions

Solving a multi-parameter inverse problem is conceptually and mathematically identical to a mono-parameter problem. Physically, however, the complications are increased. Consider the problem of solving simultaneously for the P-wave velocity (Vp) and the density (ρ) as a function of depth. It will become extremely difficult to define a resolution width. Any resolution width will imply a coupling between different physical parameters (Vp and ρ) at different depths. The interpretation of the estimated model is very delicate in such a situation. The only acceptable resolution would be a Dirac-like shape, which will result in a big error in the final model. In multi-parameter problems it seems very difficult to obtain an optimal trade-off between error and resolution. Similarly, the choice of C_m in the generalized least squares is a major problem. Already we don't know very well the error on a single physical parameter, its correlation with other physical parameters is even more difficult to estimate. Even if we know that in an elastic Earth, the velocity is related to the density via the Lame constants at a certain depth, the Lame constants themselves are not known so that such a theoretical relationship is of no great help.

Another interesting problem is coming from inconsistency. In a problem involving different physical parameters, there are several equivalent ways of formulating the problem. For instance choosing [Vp,Vs] or [Vp,Vp/Vs] should give comparable results for a given inverse operator. Unfortunately, very often, the same algorithm applied to different parameter sets and the same data give incompatible results, and the inverse operator is said to be inconsistent. The problem comes from the a priori model covariance which is implicit in some inverse operators. Remember that C_m defines the metric in (11) with which the model variation is weighted. Take for instance the Backus-Gilbert inversion. It assumes and implicit model covariance $C_m=\infty I$ (Tarantola, 1987). If we change the parameterization from m to m' ($m'=Pm$), and if we apply Backus-Gilbert again, the inversion assumes now $C_{m'}=\infty I$. The two different norms defined by the metrics C_m^{-1} and $C_{m'}^{-1}$ used in the inversion are incompatible. Covariances change according to certain rules when changing the parameters, and the right covariance to be used is $C_{m'}=PC_mP^t$ which is generally different from ∞I. The compatible $C_{m'}$ however cannot be naturally introduced in the Backus-Gilbert scheme.

A practical example has been given by Leveque and Cara (1985) who tried to study the anisotropic parameter $\xi=V_{SH}^2/V_{SV}^2$ in the upper mantle of the Pacific Ocean. In two different inversions of 28 phase velocity data, they have chosen the physically equivalent parameterizations of the problem $[V_{SH},V_{SV}]$ and $[\xi,V_{SV}]$. They used Wiggins' (1972) algorithm and the same eigenvalue cut-off level in the two inversions. The two models are very different, making the geophysical interpretation strongly dependent on the choice of parameters (fig. 3). Again the explanation is that Wiggins' algorithm uses an implicit diagonal model covariance, which is not compatible in the two parameterizations.

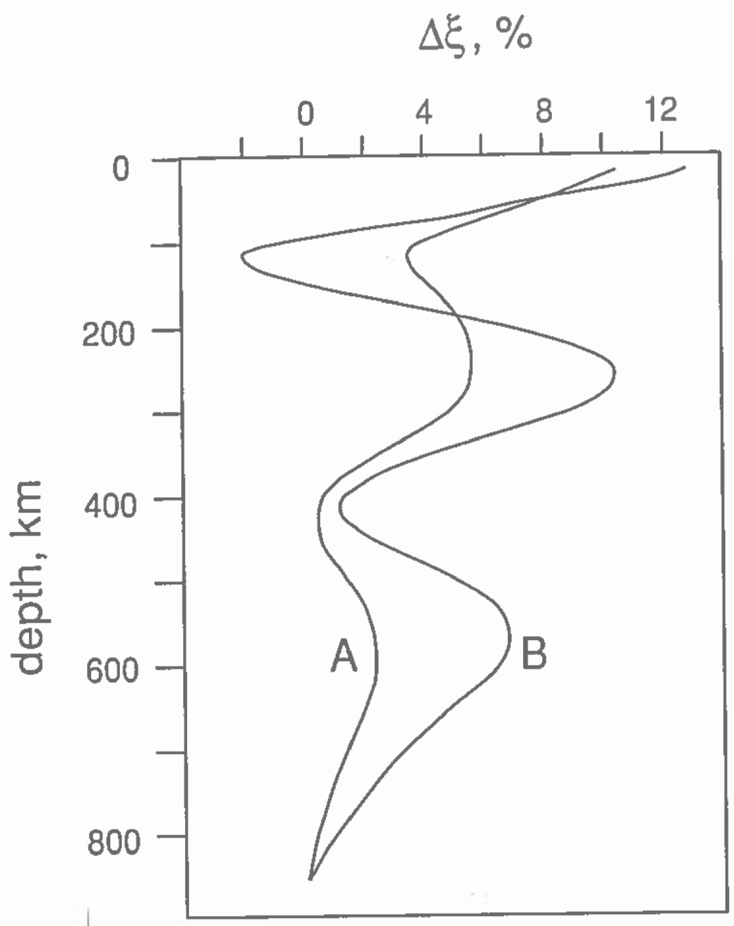


Figure 3: Example of the effect of inconsistent implicit a priori information: anisotropy in the upper mantle obtained from two different parameterizations, after Leveque and Cara, 1985.

Due to the coupling (or trade-off) between physically different parameters and due to the inconsistency problem, the choice of the set of parameters is of primary importance when inverting multi-parameter systems. The comparison of two different inversion results is not straightforward and this might be the main reason for differences between different models published so far in seismological literature. Both problems show the importance of the a priori model covariance. When using an algorithm with implicit model covariance, it seems

unavoidable to normalize the problem in a proper way before inversion. Obviously, a lot of research has to go into the correct estimation of C_m.

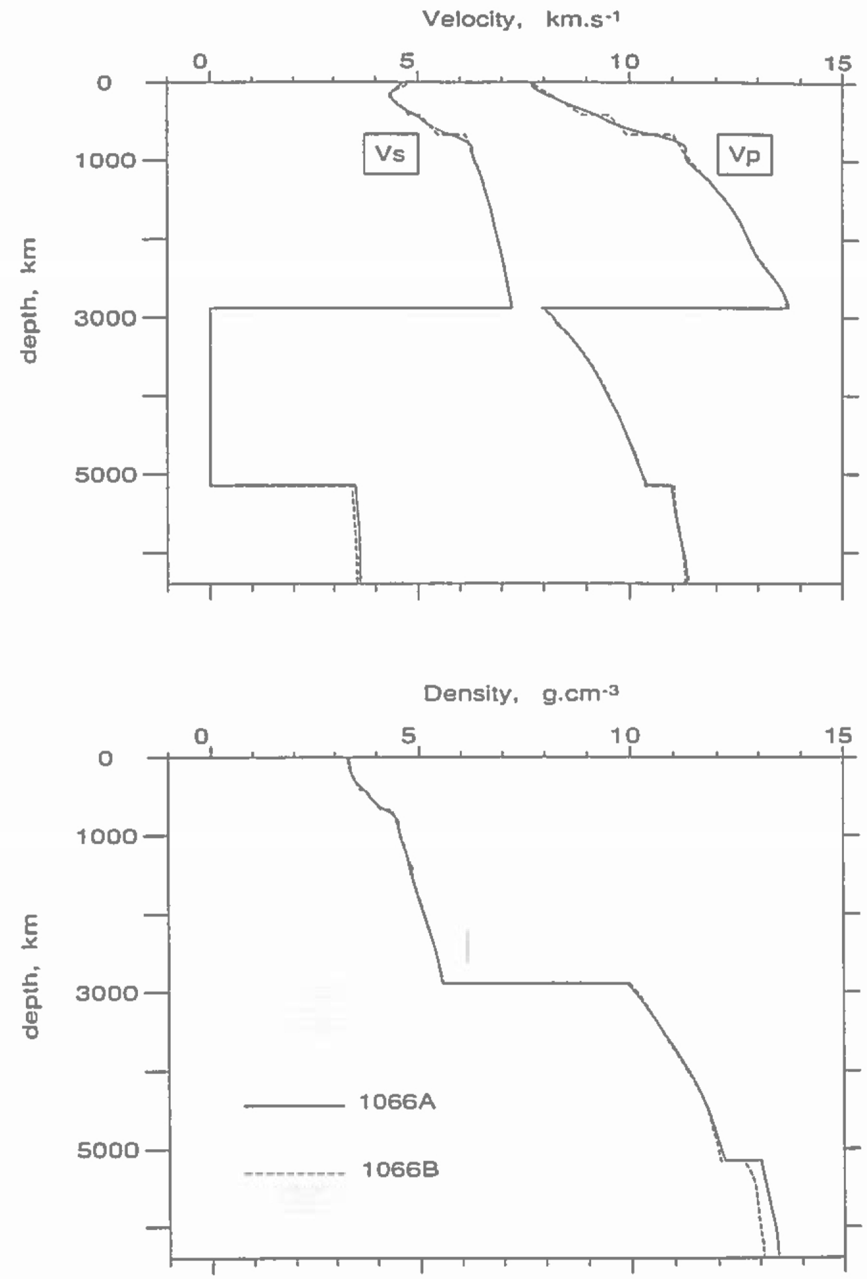


Figure 4: Earth models after Gilbert and Dziewonski, 1975.

3. RECENT APPLICATIONS IN SEISMOLOGY

3.1. Backus-Gilbert inversion

A first example is taken from Gilbert and Dziewonski (1975). They used 1066 independent normal mode data to solve a double inverse problem. In a first step, they solve for the Earth's structure and in a second step for the source mecanism of the different earthquakes. The two problems are not independent, but they decoupled them and developed an iterative method of processing the observed data. They used the Backus-Gilbert method to fit the data in their error bars (Gilbert, 1971). We show (fig.4) the two Earth models 1066A and 1066B obtained with the same algorithm but for two different starting models (one without and one with discontinuities). These discontinuities are also the essential differences in the final models.

This work has set the start for a general search in better Earth models, and emphasized the importance of a common database for different studies, which can then more easily be compared.

3.2. Seismic tomography

Seismic tomography is after hypocenter determinations the most common inverse problem in seismology. The data are generally travel time residuals, the difference between the observed travel times and the theoretical ones read in travel time tables. The residuals are assigned to velocity heterogeneities around a laterally homogeneous reference model. The problem is entirely linear and can be solved using any of the previously discussed algorithms. Figure 5 shows the result of a project in the upper Rhinegraben (Achauer et al.1989). The cross section indicates a well marked low velocity graben and high velocity mountain roots in the Vosges mountains and the Black Forest.

In general, the correlation of tomographic images with surface tectonic features is well established, while the interpretation of deep heterogeneities is more difficult.

3.3. Waveform modelling of surface waves

The retrieving of the Earth's mantle structure from the waveform of long-period Rayleigh waves leads to another interesting inverse problem. In the example presented here (Leveque et al., 1991), the highly non-linear dependency between data and parameters is avoided by a reparameterization of the problem. The information contained in the seismogram is concentrated into a few "secondary observables" by filtering and other signal processing methods. These new observables are designed to have a more linear behaviour with respect to the model parameters. An actual seismogram of at least 128 data points can be compacted into 21 secondary observables, leading to a very simple inverse problem. It can be seen on figure 6 that very few information is lost, since the synthetic seismogram obtained from the final model is very close to the recorded one.

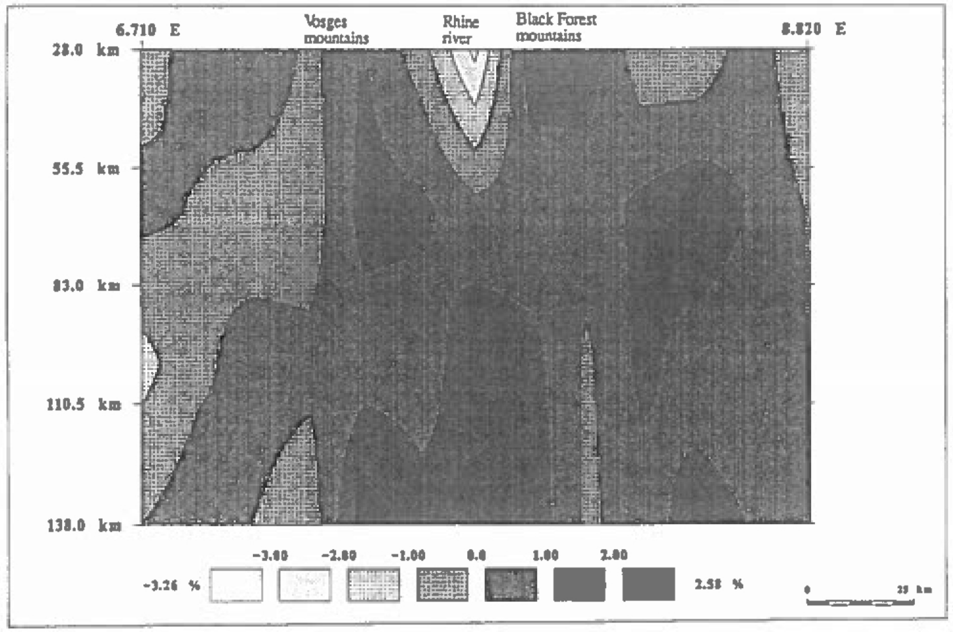


Figure 5: East-West cross-section of a 3-D P-wave velocity perturbation model through the upper Rhinegraben.

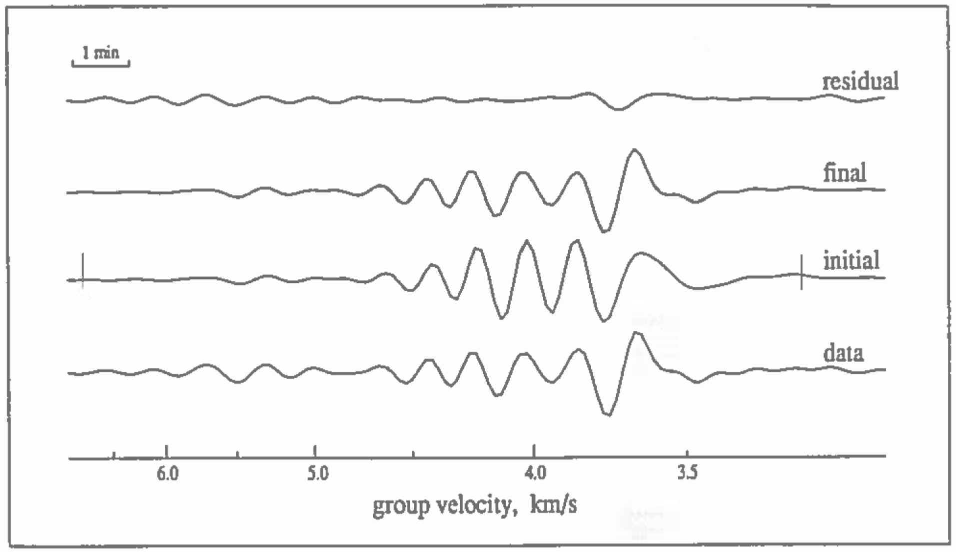
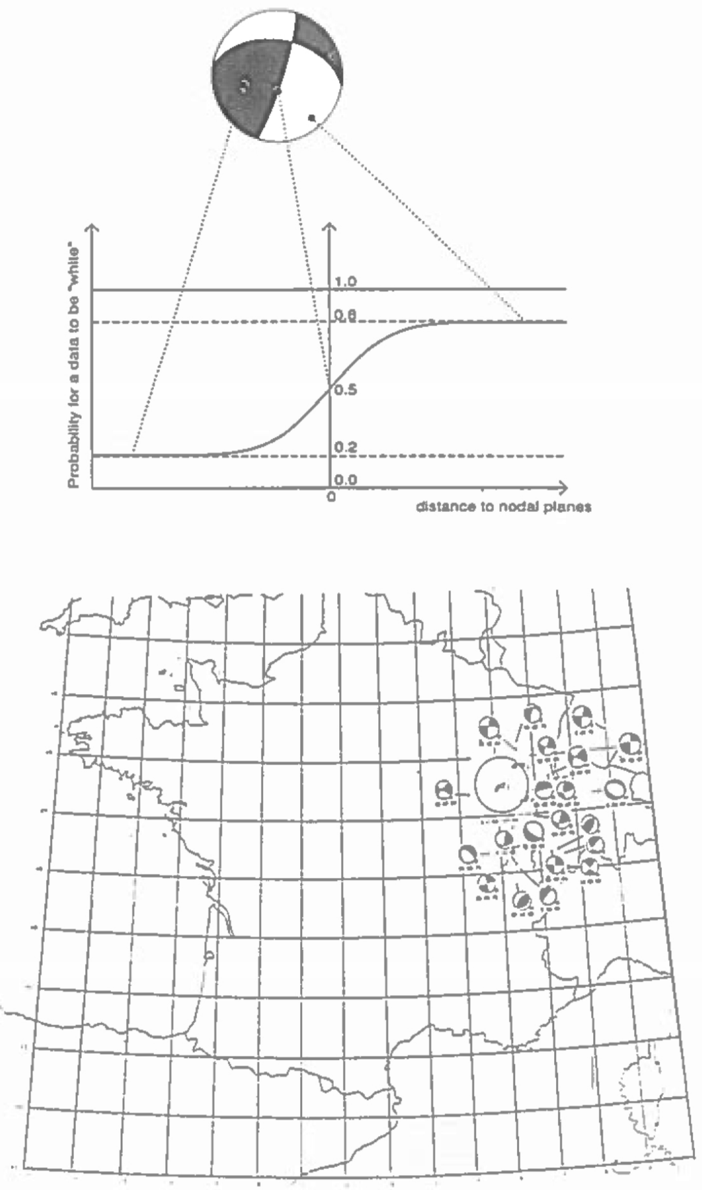


Figure 6: Waveform inversion of the 13 September 1986 Kermadec event, after Leveque et al., 1991.

a



b

Figure 7: a) Loose theory describing the polarity of seismic waves as a function of distance from the nodal plane, after Rivera, 1990.
b) Regional stress tensor in Eastern France inferred from wave polarities, after Delouis, 1988.

This example illustrates how a well-chosen reparameterization can simplify the inverse problem, even if the initial physical problem is very difficult to handle. The remaining question is how to find such well-behaved secondary observables? Unfortunately, there is no general answer, as far as we know.

3.4. Inexact theory

Due to unknowns in the physical modelling of the problem, it is sometimes better to use a probabilistic theory instead of a deterministic one. An illustration of such a situation is shown on figure 7a, where the polarity of the first seismic wave is represented as a function of the geographic position of the recording station (compressions are shown in black and dilatations in white) (Rivera, 1990). For stations located far away from a nodal plane of the seismic radiation (black-white border), there is little doubt about the theoretical polarity: the wave is clearly a compression or a dilatation, and we only have to account for measurement noise, fixed to a probability of 0.2 here. When the station moves closer to a nodal plane, the theoretical amplitude decreases and the importance of the unknown source process increases. For a station located on a nodal plane, the probability for the theoretical data to be "black" or "white" is 0.5.

This kind of loose theory is easily accepted in Tarantola and Valette's (1982) formulation of the inverse problem, and has been used to process polarity diagrams from many earthquakes in France in order to obtain a regional stress tensor compatible with the whole set of data (Delouis, 1988). This stress tensor is represented (figure 7b) by its 3 eigendirections $\sigma 1$, $\sigma 2$ and $\sigma 3$, and is compatible with the NW-SE push of the African plate against the European plate in this region.

4. CURRENT RESEARCH TOPICS

The philosophy of the linear (or linearized) inverse problem is now very well understood, and seismologists have all the necessary tools to solve it and make an interpretation of the final model. As pointed out before, the chosen parameterization of the problem is of great importance. It seems very promising to look for a parameter set which concentrates the same amount of information in fewer data (Trampert, 1990; Leveque et al., 1991). This allows to take into account a greater number of information without drastically increasing the dimensions of the problem. Related to the parameterization is the choice of the a priori information. A lot of effort goes into the search of how to introduce a priori information from different geophysical fields in the data (Nataf, 1986). For instance, how could we constrain,prior to the inversion, the core-mantle boundary topography.

The non-linear approach has essentially been attempted with Monte-Carlo techniques (Keilis-Borok and Yanovskaya, 1967; Press, 1968), but we are basically at the point to try to understand how non-linear inversions work (e.g. Snieder, 1990). An interesting thing to notice is that it is not the problem itself which is linear or not, but its parameterization. The idea is then to find a parameterization which is as linear as possible with respect to the model.

If you are interested in the elastic quality factor for instance, you know that the variation in amplitude of a wave is exponentially related to Q. Using the logarithm of the amplitude instead of the amplitude itself will naturally linearize the problem. Here we join the previous discussion of the importance of a good parameterization.

Finally, we should drop the ad hoc separation of the problem into source and structure. A more general formulation seems necessary to analyse the importance of the trade-off between source and structure. This needs a complete waveform tomography, and even if we manage to concentrate the information, we still have to invert very large problems. It is thus necessary to try to improve the algorithms in terms of speed and memory consumption. Most algorithms used in seismology directly solve (4), (5), (6) or (12) in matrix form. It is easily seen that this needs a considerable amount of computer memory, so that we are soon limited in the size of the problem which can be treated. Interesting alternatives are the row-action and projection methods (e.g. Van der Sluis and Van der Vorst, 1987). The row-action methods act on the matrix G as it stands without making any change to it. They work on one row of G at a time, hence their name, and update the model iteratively. The most popular method of this type in seismology is the Simultaneous Iterative Reconstruction Technique (SIRT) designed by Gilbert (1972) for medical tomography. It handles elegantly and simply a very large system, but has the drawback of introducing implicit a priori information depending of the theory matrix G itself. Trampert and Leveque (1990) showed how to correct this implicit behaviour and changed the original algorithm so that it becomes a special case of (12). With a proper normalization, using covariance operators, before inversion, it is possible to solve any linearized problem with SIRT introducing explicitly any m_0, C_m and C_d in the solution. Other interesting approaches are the projection methods. The system is projected into the Krylov sub-space and solved iteratively without explicitly computing the projection. The Least Squares Conjugate Gradient method (LSQR) developped by Paigne and Saunders (1982) seems to be particularly fast and efficient (Nolet, 1985). The solution has the form of (5), and, as it is not a special case of the generalized least squares, it is not obvious how to normalize the problem so that you can explicitly take into account any a priori information.

ACKNOWLEDGMENTS

We thank Alfred Muller for providing figure 1 and Michel Granet for providing figure 5.

REFERENCES

Achauer U Glahn A Granet M Wittlinger G and Slack P D 1989 *EOS* 70 1221 (Abstract)
Aki K and Richards P G 1980 *Quantitative Seismology* (San Fransisco: Freeman)
Backus G and Gilbert F 1967 *Geophys. J. R. astr. Soc.* 13 247-276
Backus G and Gilbert F 1968 *Geophys. J. R. astr. Soc.* 16 169-205

Backus G and Gilbert F 1970 *Phil. Trans. R. Soc. Lond.* A266 123-192
Cara M and Leveque J-J 1987 *Phys. Earth Planet. Int.* 47 246-252
Das S and Kostrov B V 1990 *J. Geophys. Res.* 95 6899-6913
Delouis B 1988 *DEA thesis* (Strasbourg: University Louis Pasteur)
Deschamps A Lyon-Caen H and Madariaga R 1980 *Annal. Geophys.* 36 179-190
Dziewonski A M 1984 *J. Geophys. Res.* 89 5929-5952
Franklin J N 1970 *J. Math. Anal. Appl.* 31 682-716
Gilbert P 1970 *J. theor. Biol.* 36 105-117
Gilbert F 1971 *Geophys. J. R. astr. Soc.* 23 925-928
Gilbert F and Dziewonski A M 1975 *Phil. Trans. R. Soc. Lond.* A278 187-269
Herglotz G 1907 *Physikalische Zeitschrift* 8 145-147
Jackson D D 1979 *Geophys. J. R. astr. Soc.* 57 137-157
Jeffreys H 1939 *Mon. Not. R. astr. Soc. Geophys. Supp.* 4 498-533
Kanamori H and Cipar J J 1974 *Phys. Earth Planet. Int.* 9 128-136
Keilis-Borok V J and Yanovskaya T B 1967 *Geophys. J. R astr. Soc.* 13 223-234
Lanczos C 1961 *Linear Differential Operators* (London: Van Nostrand)
Langston C A and Helmberger D V 1975 *Geophys. J. R. astr. Soc.* 42 117-130
Levenberg K 1944 *Quaterly of appl. math.* 2 164-168
Leveque J-J and Cara M 1985 *Geophys. J. R. astr. Soc.* 83 753-773
Leveque J-J Cara M and Rouland D 1991 *Geophys. J. Int.* 104 433-720
Matsu'ura M Hirata N 1982 *J. Phys. Earth* 30 451-468
Montagner J-P and Nataf H-C 1988 *Geophys. J.* 94 295-307
Nataf H-C 1986 *J. Geophys. Res.* 91 7261-7307
Nolet G 1985 *J. Comp. Phys.* 61 463-482
Nolet G 1987 *Seismic Tomography* ed G Nolet (Dordrecht: Reidel Publishing Company) pp 301-322
Paigne C C and Saunders M A 1982 *ACM Trans. Math. Softw.* 8 43-71
Penrose R 1955 *Proc. Cambridge Phil.Soc.* 51 406-413
Press F 1968 *J. Geophys. Res.* 73 5223-5234
Rivera L 1990 *Ph.D. thesis* (Strasbourg: University Louis Pasteur)
Snieder R 1990 *Geophys. J. Int.* 101 545-556
Tanimoto T 1987 *Geophys. J. R. astr. Soc.* 89 713-740
Tarantola A 1987 *Inverse Problem Theory* (Amsterdam: Elsevier)
Tarantola A and Valette B 1982a *J. Geophys.* 50 159-170
Tarantola A and Valette B 1982b *Rev. Geophys. Space Phys.* 20 219-232
Trampert J 1990 *Ph.D. thesis* (Strasbourg: University Louis Pasteur)
Trampert J and Leveque J-J 1990 *J. Geophys. Res.* 95 12553-12559
Van der Sluis A and Van der Vorst H A 1987 *Seismic Tomography* ed G Nolet (Dordrecht: Reidel Publishing Company) pp 49-83
Wiechert E 1907 *Nach. Ges. Wiss. Goettingen* Math. Phys. Klasse 415-533
Wiggins R A 1972 *Rev. Geophys. Space Phys.* 10 251-285
Woodhouse J H and Dziewonski A M 1984 *J. Geophys. Res.* 89 5953-5986
Woodhouse J H Giardini D and Li X D 1986 *Geophys. Res. Lett.* 13 1549-1552