Physics of the Earth and Planetary Interiors, 10 (1975) 292–298

ON THE INVERSION OF GLOBAL ELECTROMAGNETIC INDUCTION DATA

R.S. ANDERSSEN

Computer Centre, Australian National University, Canberra, A.C.T. (Australia)

(Received May 12, 1975)

The initial phase of any inversion of geophysical data must examine the question of the existence of globally distinct solutions. Previous inversion studies for global electromagnetic induction (GEMI) data are reviewed from this point of view. A basic inversion strategy for geophysical data is considered. It is concluded that future progress depends on the use of synthetic data to resolve questions about the potential constraining power of GEMI data.

1. Introduction

In the general field of geophysical data inversion, a number of independent methods such as Monte Carlo inversion (Press, 1970; Anderssen et al., 1972), linearization (Backus and Gilbert, 1967, 1970), hedgehogging (Keilis-Borok and Yanovskaya, 1967), and edgehogging (Jackson, 1973) have been applied, but little attention has been paid to what the general philosophy underlying data inversion should be. We therefore aim to examine this question here, at least from the point of view of inverting GEMI data, and thereby generate a basis for the comparison of methods proposed to date.

1.1. Direct inversion through the use of an inversion formula

Though an inversion formula is known for frequency response GEMI data, it does not yet have a satisfactory computational implementation, because knowledge of the whole frequency response spectrum is required for one of the spherical harmonic modes (see Bailey, 1973, p. 238). The available frequency response data are in fact very limited, because of the nature of the observational problem involved. While the known inversion formula for frequency response GEMI data cannot be used now, it may have theoretical importance. It may illuminate the underlying non-uniqueness problem and the extent to which different types of data constrain solutions. Though some such aspects have been examined (see, for example, Bailey, 1970, 1973; and Weidelt, 1972), full advantage of its potentialities does not appear to have been exercised. For example, no use of it has been made to investigate which sections of the frequency response spectrum exert the strongest constraints.

1.2. Linearization of the problem formulation

In recent years, linearization of the problem formulation has played an important role as a tool in geophysical data inversion. Much of its popularity is due to the work of Backus and Gilbert (1967, 1968, 1970), who have formalized it as a tool in geophysics, though the essence of the technique dates back to Newton (see, for example, Rall (1969, chapter 4); and Vainberg (1964, chapter 8)), and the Backus-Gilbert framework is quite general. In the mathematical literature, it is often referred to as the Newton-Kantorovich method (see, for example, Mikhlin and Smolitskiy, 1967). In a geophysical context, many authors have stressed the essential limitations of the method (for examples from the geophysical literature see Jackson, 1973, introduction; and Backus and Gilbert, 1970, section A.1), but none, except for Sabatier (1974, section 4), have actually illustrated explicitly the pitfalls which surround its use.

In Section 2, we aim to show that:

(A) The use of linearization can lead to the introduction of assumptions which are not compatible with the assumptions which underlie the problem formulation.

(B) When using a linearization method, there is no guarantee, unless the contrary is established, that, for a given starting solution, the method:

(i) Will not oscillate (without converging), or

(ii) Will converge to the solution which is closest to the starting solution.

1.3. Direct numerical methods of solution

When using the word linearization, it is necessary to distinguish between the linearization of the problem formulation and the use of linearized numerical methods to help solve the non-linear formulation by some direct procedure. Sometimes they yield the same computational procedure, but often they don't. (For example, compare the linearization method of Parker (1971), which neglects the quadratic term in a frequency response version of the Lahiri and Price (1939) equations, with the direct methods examined by Anderssen (1968), where no terms are neglected because the Lahiri and Price (1939) equations are linear when the conductivity distribution is known.) Both have their problems, but the former is more invidious than the latter when it yields a formulation which does not reflect the character of the non-linear formulation from which it has been derived (see Sabatier, 1974, section 4).

When inverting data, it must be explicitly assumed that unless the contrary is established, given data support globally distinct ("G-far" in the Backus and Gilbert (1970) terminology) solutions.

The interpretation of what is meant by a globally distinct solution is problem dependent. For example, the definition of "G-far" solutions of Backus and Gilbert (1970) is qualitative rather than quantitative. For this reason, we shall regard two solutions as being globally distinct, if a geophysical explanation of each must appeal to basically different geophysical, physical, chemical and geochemical concepts. For example, we should regard as globally distinct two solutions which are identical except that one has a large sharp discontinuity while the other does not. In the context of free oscillation data inversion, Anderssen and Cleary (1974) have shown that the geophysical behaviour and interpretation of Earth models depends heavily on the size, sharpness and position of the discontinuities which they contain.

Two different computational strategies have been developed for the inversion of data. They are:

(1) *Non-uniqueness modelling.* For given data, determine the degree of non-uniqueness which they support within given a priori bounds for the solutions.

(2) *Refinement modelling.* For given data and a sufficiently accurate approximation to what is "thought" to be a "correct" solution supported by the data, determine a better approximation to this solution.

Though the use of (1), the possible existence of globally distinct solutions to a given inversion problem can be examined. In situations where the degree of non-uniqueness (i.e., extent of the variation among the globally distinct solutions), and hence the data, does not support the existence of such solutions, we obtain justification for the assumption of (2). If (2) is applied when globally distinct solutions exist, then an approximation to one of them is all that is obtained.

2. Linearization and a basic inversion strategy

We start by examining the effect of the linearization of a problem formulation. For clarity, we examine a rather simple problem, since justification for the conclusions reached (see (A) and (B) of the Introduction) in the broader context of geophysical data inversion can be found in Sabatier (1974, section 4).

For the problem of determining a zero of a one-dimensional non-linear equation $f(x) = 0$, the standard linearization procedure is Newton's method: for a given starting solution x_0, evaluate the iteration:

$$x_{n+1} = x_n - f(x_n)/f'(x_n), \qquad n = 0, 1, 2, \ldots$$

The theory which underlies its use, just like the results of Backus and Gilbert (1970), for the linearization of geophysical inverse problems, asserts that Newton's method converges, if:

(a) x_0 is sufficiently close to the required zero, and

(b) $f(x)$ is sufficiently well behaved in a neighbourhood containing both the starting solution and the zero. If either fails to hold, then Newton's method will fail.

To illustrate what can happen when either (a) or (b) fails, we consider two examples.

Example 1. Consider the situation where:

$f(x) = x^4 - 6x^2 - 11$

and the starting solution x_0 is given the values 0.5, 0.75, 1.25, 2.0. The results are shown in Table I. Since the zeros of $f(x) = 0$ are $\pm(3 + 2\sqrt{5})^{1/2}$, we can draw the following conclusions. There exist starting solutions such that Newton's method oscillates without converging, converges to the zero which is not the closest to the starting solution, and behaves normally. In the more general framework of geophysical inversion,the first illustrates that linearization procedures can oscillate without converging, while the second shows that there is no guarantee that a linearization procedure will converge to the closest solution, and that the convergence may be slow and via intermediate-solutions which are very poor approximations to any of the possible solutions.

TABLE I

Newton sequences $\{x_n; n = 0, 1, \ldots, 16\}$ for the solution of $f(x) = x^4 - 6x^2 - 11$

n	x_n	x_n	x_n	x_n
0	0.5	0.75	1.2500	2.000
1	−17.614	−1.1725	−1.2451	4.3750
2	−29.468	1.1049	1.2338	3.5235
3	−22.127	−1.0360	−1.2092	3.0062
4	−16.629	1.004	1.1617	2.7796
5	−12.518	−1.0	−1.0909	2.7351
6	−9.4512	1.0	1.0267	2.7335
7	−7.1739	−1.0	−1.0022	2.7335
8	−5.4993	1.0	1.0	2.7335
9	−4.2943	−1.0	−1.0	2.7335
10	−3.4708	1;0	1.0	2.7335
11	−2.9784	−1.0	−1.0	2.7335
12	−2.7716	1.0	1.0	2.7335
13	−2.7346	−1.0	−1.0	2.7335
14	−2.7335	1.0	1.0	2.7335
15	−2.7335	−1.0	−1.0	2.7335
16	−2.7335	1.0	1.0	2.7335

Example 2. For the odd function:

$f(x) = x^p \exp(-x^2), \quad f(-x) = -f(x), \quad p > 0$

which has only one zero at $x = 0$, it is not difficult to establish the following results:

(i) For $p > \frac{1}{2}$, Newton's method fails to converge, if the starting solution lies in either of the intervals $(-\infty, -[(2p-1)/2]^{1/2})$ or $[(2p-1)/2]^{1/2}, \infty)$.

(ii) For $p \leqslant \frac{1}{2}$, Newton's method fails to converge no matter how close to the origin the starting solution is placed.

The first results shows that, even for reasonable looking starting solutions (points well on the origin side of the maximum and minimum of $f(x)$), Newtons's method can fail to converge. The second shows that there exist quite reasonable and simple functions for which Newton's method fails, because their behaviour in the neighbourhood of a zero is pathological. For an example of a system of equations for which the convergence of Newton's method does not hold, see Rall (1969, p. 190).

Situations of the nature cited in these two examples have occurred in the application of the procedure of Backus and Gilbert (1967, 1970) to geophysical data inversion, but the possibility that it is a consequence of linearization has been overlooked. For example, Parker (1971, p. 130), concludes that the lack of convergence of his linearization method for the inverse problem of electrical conductivity is due either to "the starting model always being too far from the true structure*, or no Earth model exists which is compatible with the data". He neither acknowledges nor rules out the possibility that the lack of convergence is a consequence of the linearization used. Even though this may not be the case (see, for example. Jady, 1974b), it is a fact which cannot be ignored, as the above examples show.

For the simple case of Newton's method treated above, the validity of the linearization follows from the validity of (a) and (b). For more complex problems, such as arise in geophysical inversion, the validity of the linearization may fail to hold or will only hold under additional assumptions (see, for example, Sabatier, 1974, section 4).

On the basis of the above, we can conclude that linearization is only applicable, if an approximate solution is known which is sufficiently close to a possible solution and if the resulting linearization procedure is valid. Thus, the stage at which lineariza-

* Really a possible structure, since the data supports either no or infinitely many solutions, but not the true structure.

tion is applicable in a general modelling strategy is when the local structure of the formulation is known; i.e., at the refinement modelling stage.

However, we can only obtain the required local structure information in one of two ways:

(1) Assume that we can determine this local information from independent geophysical and geochemical results. Since the aim of geophysical data inversion (especially, GEMI-data inversion) is usually aimed at obtaining independent information which can be used to clarify what the independent geophysical results are, this approach should be avoided because of its potential circularity.

(2) Determine whether or not the data supports globally distinct solutions.

Thus, in an inversion, the initial aim must be the determination of whether or not the data support globally distinct solutions, for depending on the answer, the inversion must proceed differently. In fact, we have the following possibilities:

(i) If there appear to be no solutions, then there is an inconsistency within the data which must be removed.

(ii) If there appear to be no globally distinct solutions, then a model refinement procedure could be applied, along with additional assumptions to test additional hypotheses about the structure of models.

(iii) If there appear to be only two or three globally distinct solutions, then we know that the data do exert a strong constraining influence, and we can seek independent geophysical evidence or data to reduce, if possible, the number of globally distinct solutions to one.

(iv) If there are many globally distinct solutions, then we know that the data do not constrain the models sufficiently strongly, and therefore attention has to concentrate on improving the accuracy and/or the determination of new data which impose new independent constraints on models. In this case, it is also necessary to check whether the nature of the data is such than only qualitative information is resolvable.

We therefore conclude that the two steps of a Basic Inversion Strategy for the inversion of (geophyscial) data must be:

(a) Test for the existence of globally distinct solutions. If this is not done before a model refinement procedure is applied or a specific solution is determined as the model which best fits the data, then it has been tacitly assumed that either no globally distinct solutions exist, which may be a false assumption, or that the solution generated is only the one which best fits the data from a particular set of all the possibilities.

(b) Proceed along the lines (i), (ii), (iii) and (iv) laid down above, depending on the number of globally distinct solutions discovered as a result of the implementation of (a).

This inversion strategy represents a basis for assessing proposed methods for GEMI-data inversion and will be the one used in this paper.

3. A survey of GEMI-data inversion studies

In the light of the proceeding discussion, we now classify the different types of GEMI-data inversion studies proposed since 1955.

Our general conclusion must be that none has adequately examined the extent of the underlying non-uniqueness or made a systematic search for globally distinct solutions. In one way or another, they have been concerned with model refinement rather than non-uniqueness modelling. Nevertheless, non-uniqueness has been discussed. Anderssen (1968) has shown that greatly different models can fit the data equally well. Banks (1969) derived upper and lower bounds for his "best-fitting" model, and showed that they were extremely broad. Using synthetic data, Anderssen (1970a) has shown that the non-uniqueness is an inherent feature of the GEMI-data inversion problem (due to the inadequate nature of experimental data) and not solely a feature of inaccurate data. In addition, the quantized nature (i.e., sets of solutions of basically different structure fit the data with the same degree of precision) of the non-uniqueness was established in that paper. Finally, using a linearization of the GEMI equations, Parker (1971) confirmed Bank's findings.

We classify the GEMI-data inversion studies since 1955 as:

(1) *The electromagnetic frequency response method.* Several applications of this method for variations at discrete frequencies have been made, and reviewed recently by Price (1970). It was not until the work of Banks (1969) that the full potentialities of the method were implemented. Linearization of the GEMI equa-

tions has only been applied for this class of methods.

(2) *The direct solution of the Lahiri and Price equations.* In such methods, the time-domain external and internal components of a magnetic disturbance are used. By assuming a known structure (model) for the conductivity, the Lahiri and Price (1939) equations can be treated as a linear parabolic partial differential equation and solved by standard numerical methods, such as finite difference methods, to determine the internal field at the surface of the Earth corresponding to a given external field and conductivity model. The aim is to adjust the conductivity model until the calculated internal field at the surface approximates the observed suitably closely. Except for the odd publication (see, for example, Price (1970) and Anderssen (1968)) little use has been made of this method.

(3) *Bailey's causality method.* This is an iterative method which is based on the uniqueness theorem of Bailey (1970).

(4) *Weidelt's implementation of the Gelfand-Levitan method.* This is the integral equation method which arises when the Gelfand-Levitan method for Schrödinger's equation is applied to the GEMI equations (Weidelt, 1972).

Since, as Bailey (1973) points out, both he and Weidelt have found that their methods do not perform as well as modelling methods when applied to noisy and truncated data, the last two methods must be regarded as unsatisfacory at this point in time. However, this does not rule them out as potential methods of the future. The recently published work of Jady (1974a) gives some indication of the potentialities of such methods.

The above list does not exhaust the methods which could be applied. In particular, there exist two methods which have not been implemented for GEMI data, but which could be, and therefore deserve a mention. They are:

(5) *Regularization.* (See, for example, Lavrentiev (1967).) In the implementation of this method, an appropriate stabilization condition would be used to restrict the classes of globally similar solutions around the globally distinct ones to subclasses of smooth ones. For example, we could reformulate the direct solution of the GEMI equations as a non-linear programming problem: for given a_i, $i = 0, 1, 2, ..., m$, with $a_m \neq 0$:

$$\min_{\kappa \epsilon W} \sum_{i=1}^{m} a_i \|\kappa^{(i)}\| \, , \; \kappa^{(i)} = \partial^i \kappa / \partial x^i$$

subject to the equality constraints that κ (conductivity) satisfy the GEMI equations, where W is some class of smooth functions (e.g., $W = C^{(m)}$ [0,1] the class of functions defined on the interval [0,1] with continuous mth derivatives). Other non-linear programming formulations are possible.

(6) *Monte Carlo inversion.* This method is specifically designed to search for globally distinct solutions which lie between a priori non-uniqueness bounds defining the extent to which geophysically realistic solutions can vary. A method for determining such bounds for GEMI-data inversion has been discussed by Anderssen (1970b). Any one of the methods (1), without linearization, (2) or (5) can be used to implement it along the lines discussed in Anderssen and Seneta (1971, 1972), and Anderssen et al. (1972).

The advantage of (6), from the point of view of the inversion strategy of section 2, is that it exhibits the quantization present within the non-unique solutions which satisfy the data and allows hypotheses about the number of globally distinct solutions supported by the data to be tested. It avoids the averaging inherent in the linearized estimates of the type derived by Parker (1971).

Because the GEMI equations (viz., the Lahiri and Price (1939) equations) *define a linear formulation* for relating the internal to the external field at the surface of the Earth once a model for conductivity is specified, there does not appear to be a strong case for using linearization methods of the type proposed by Parker (1971) and Bailey (1973) for refinement modelling, since both neglect terms which it is not necessary to neglect. It can be argued that linearization is only justifiable when the nature of the mathematical formulation to be solved: (i) cannot be reformulated as a linear problem without neglecting terms; (ii) methods for its solution are not known; and/or (iii) the computing time required to solve it without the use of linearization is prohibitive. None of these conditions apply for the refinement modelling of GEMI data.

On the other hand, a case can be argued for the use of linearization methods to explore the qualitative dependence of conductivity as a function of depth on the observed external and internal fields at the surface of the Earth. However, because of the pitfalls associated with the use of linearization, the viability of linearization in a particular application should first be tested with synthetic data, and only used to derive results from real data when such viability has been established.

The question of whether (1) without linearization, is preferable to (2), when used for model refinement, appears to be open. It is clear that (1) has the potential to yield a profile of $\kappa(x)$ to a greater depth than is obtainable from (2), whereas (2) has the potential, given appropriate quality D_{st} data, to resolve more detail of $\kappa(x)$ in the upper mantle than (1).

4. The degree of constraint of available GEMI data

Although all recent inversion studies have concluded that available data is unable to constrain a unique or even globally similar structure for global conductivity models, this does not rule out the possibility that available data can support certain qualitative conclusions. In fact, all studies have tended to confirm that, globally, $\kappa(x)$ has the following qualitative structure: (i) a region of low conductivity in the surface layers of the Earth; (ii) a region below the surface layers where conductivity increases rapidly with depth, that is, a conductivity "discontinuity" below the surface layers; and (iii) a region of increasing conductivity below the "discontinuity". That the available data may be able to constrain sharper qualitative conclusions than this seems to have been overlooked by all authors, except for the possibility of Banks (1972) and Jady (1974a).

For example, it would be of great advantage for future conductivity modelling, if the periods of the disturbances which carry most information about the position, size and shape of the conductivity "discontinuity" could be delineated. That this may be a viable possibility has been illustrated by Banks (1972) in his discussion of the interpretation of responses with very small phase. It would be useful for the conductivity modeller, if sharper rules of thumb than just "the longer the period, the deeper the conductivity structure can be resolved" were available.

From the point of view of testing for and confirming that given data supports certain qualitative structure, Monte Carlo inversion is the most natural of the above methods to use.

Further testing of the constraint, imposed on conductivity by different classes of data, could be based on the use of synthetic data. This would lead to an extension of our understanding of the following types of questions:

(i) Can GEMI data ever be expected to support finer structure for conductivity models than purely qualitative results? On the basis of present results, the answer would appear to be "no" rather than "yes".

(ii) What types of GEMI data impose the greatest constraint on the type of conductivity structure thought to exist within the Earth? An analysis of synthetic data can yield answers to this question. Even though given data may support globally distinct models, *given globally distinct models will not necessarily support globally similar data.*

5. Recent progress and the future

In the two years since the First Workshop, little progress has been made with the inversion of GEMI data, except for the publication of papers from Banks (1972), Kuckes (1973) and Jady (1974a, 1974b).

Because Rikitake (1973) reviewed Banks' (1972) paper at length at the First Workshop, we shall not discuss it in detail but only pause to note that, until data of the type foreshadowed by Banks (1972) is forthcoming, the future progress must be based on the analysis of synthetic data along the lines cited above. This at least may resolve some of the questions concerning the interpretation of low phase response data.

Kuckes (1973) showed that, when geomagnetic disturbances have a lateral scale distance which is somewhat larger than the depth to which it is necessary to probe, a systematic technique for locally analysing data and subsequently making conductivity maps of the Earth can be developed. This is an aspect of electromagnetic induction data inversion which has not been examined in this review, but is one which has considerable importance from the point of view of the study of lateral inhomogenities within the Earth.

In his paper, Jady (1974a) shows how single periodic variations alone can be used to construct uniformly conducting thick shell models which surround a perfectly conducting sphere. This modelling procedure is then used to obtain models from *Sq* variations, the 27-day variation and its harmonics, and the annual variation as well as the maximum screening effect of the oceans. Jady (1974b) uses the same modelling procedure to reanalyse the magnetic response data of Banks (1969, 1972), and thereby confirms the conclusion of Parker (1972) that this data is inconsistent.

It may be that GEMI data can only constrain qualitative type conclusions about conductivity structure in the upper mantle, at least until greatly improved data becomes available. Some attempt should be made to clarify this, as well as other questions raised in this review, through the use of synthetic data. Until this is done, the directions of future progress are left in doubt.

References

Anderssen, R.S., 1968. J. Geophys. Res., 73: 6535–6543.
Anderssen, R.S., 1970a. Pure Appl. Geophys., 80: 238–259.
Anderssen, R.S., 1970b. Earth Planet. Sci. Lett., 7: 314–316.
Anderssen, R.S. and Cleary. J.R., 1974. Asymptotic structure in torsional free oscillations of the Earth: I. Overtone structure. Geophys. J.R. Astron. Soc., 39: 241–268.
Anderssen, R.S. and Seneta, E., 1971. Pure Appl. Geophys., 91: 5–31.
Anderssen, R.S. and Seneta, E., 1972. Pure Appl. Geophys., 96: 5–14.
Anderssen, R.S., Worthington, M.H. and Cleary, J.R., 1972. Geophys. J.R. Astron. Soc., 29: 433–444.
Backus, G.E. and Gilbert, J.F., 1967. Geophys. J.R. Astron. Soc., 13: 247–276.
Backus, G.E. and Gilbert, J.F., 1968. Geophys. J.R. Astron. Soc., 16: 169–205.
Backus, G.E. and Gilbert, J.F., 1970. Philos. Trans. R. Soc. London. Ser. A, 266: 123–192.
Bailey, R.C., 1970. Proc. R. Soc. London, 315: 185–194.
Bailey, R.C., 1973. Phys. Earth Planet. Inter., 7: 234–244.
Banks, R.J., 1969. Geophys. J.R. Astron. Soc., 17: 457–487.
Banks, R.J., 1972. J. Geomagn. Geoelectr., 24: 337–351.
Jackson, D.D., 1973. Geophys. J.R. Astron. Soc., 35: 121–136.
Jady, R.J., 1974a. Geophys. J.R. Astron. Soc., 36: 399–410.
Jady, R.J., 1974b. Geophys. J.R. Astron. Soc., 37: 447–452.
Keilis-Borok, V.I. and Yanovskaya, T.B., 1967. Geophys. J.R. Astron. Soc., 13: 223–234.
Kuckes, A.F., 1973. Geophys. J.R. Astron. Soc., 32: 119–131.
Lahiri, B.N. and Price, A.T., 1939. Philos. Trans. R. Soc. London, Ser. A, 237: 509–540.
Lavrentiev, M.M., 1967. Some Improperly Posed Problems in Mathematical Physics. Springer, Berlin.
Mikhlin, S.G. and Smolitsky, K.L., 1967. Approximate Methods for Solution of Differential and Integral Equations. American Elsevier, New York, N.Y.
Parker, R.L., 1971. Geophys. J.R. Astron. Soc., 22: 121–138.
Parker, R.L., 1972. Geophys. J.R. Astron. Soc., 29: 123–138.
Press, F., 1970. Phys. Earth Planet. Inter., 3: 3–22.
Price, A.T., 1970. Q. J. R. Astron. Soc., 11: 23–42.
Rall, L.B., 1969. Computational Solution of Non-linear Operator Equations. Wiley, New York, N.Y.
Rikitake, T., 1950. Bull. Earthquake Res. Inst., Tokyo Univ., 28: 45–100, 219–262, 263–283; 29: 61–69.
Rikitake, T., 1973. Phys. Earth Planet. Inter., 7: 245–250.
Sabatier, P.C., 1974. Proc. R. Soc. London, Ser. A, 337: 49–71.
Vainberg, M.M., 1964. Variational Methods for the Study of Non-linear Operators. Holden Day, San Francisco, Calif.
Weidelt, P., 1972. Z. Geophys., 38: 257–289.