

Agnostic geophysics

Alan Cook

Selwyn College, Cambridge, U.K.

Abstract

The physics of the Earth and the planets itself limits the knowledge that can be obtained of it. Two principle limitations are set in inverse problems and by chaotic dynamics; examples of each are given in this essay and some implications are discussed.

Key words *inverse theory – chaos*

1. Introduction

I am very gratified to have the opportunity to contribute an essay in honour of Michele Caputo for our acquaintance goes back many years to the General Assembly of the International Union of Geodesy and Geophysics in Toronto in 1957, the occasion also on which I first met Antonio Marussi with whom I was to collaborate in various ways over many years, and whose assistant Michele was at that time. Among his many studies Michele has written on the two topics which I take up here, geophysical inverse theory and chaotic dynamics.

The Toronto meeting of the IUGG coincided with the greatest development of geophysics ever to have taken place, for it was as we returned from that meeting that news came that Russia had launched the first artificial satellite; observations of it gave us within a very short time a far more secure value for the flattening of the Earth than ever before. Geophysics has never been the same since. At the next General Assembly in Helsinki the first observations of free elastic oscillations of the Earth were reported. At much the same time there came the first really new observations of

the Earth's magnetic field and ideas of its origin since the survey and writings of Edmond Halley at the end of the seventeenth century. All those technical developments greatly extended the scope of geophysics as well as offering new possibilities in existing studies, but those very extensions and possibilities showed that geophysics had its limitations. The limitations set by inverse problems were the first to be recognised, now chaotic dynamics is being seen to be relevant to many issues in geophysics.

2. The inverse problem

One problem above all characterises the physics of the Earth and the planets, the inverse problem (Tarantola, 1987; Parker, 1994); it is stated very simply. We make observations of quantities on the surface of a planet or outside it, what may we infer about the interior? In formal terms, we postulate some function of position within the planet (density for example), from which it is possible to calculate a functional as a function of position over the surface (the value of the attraction of gravity). Suppose that ξ is a function of the vector position \mathbf{r} within the planet. The functional η as a function of position Ω over the surface is

$$\eta(\Omega) = \int_V \xi(\mathbf{r}) K(\mathbf{r}, \Omega) dV$$

where $K(\mathbf{r}, \Omega)$ is the kernel that gives the ef-

Mailing address: Sir Alan Cook FRS, 8 Wootton Way, Cambridge CB3 9LX, U.K.

fect of $\xi(\mathbf{r})$ at the position Ω . When ξ is the density and η the value of gravity, $K(\mathbf{r}, \Omega)$ is D^{-2} , where D is the distance between the points \mathbf{r} and Ω .

It is immediately clear that if some $\xi_0(\mathbf{r})$ gives the observed value of η , any function $\xi'(\mathbf{r})$, where

$$\int_V \xi'(\mathbf{r}) K(\mathbf{r}, \Omega) dV = 0,$$

may be added to it and leave η unaltered.

Parker (1994, chapter 2) gives a striking example of an implausible distribution of sea floor magnetisation that would nonetheless generate the observed field.

The possible forms of null functions, η , do depend on the mathematical form of the kernel, and some have been studied in depth, especially those that arise in potential problems, but in general physical limits cannot be set from within the inverse problem itself, but only by appealing to other considerations. Thus if $\xi'(\mathbf{r})$ is the mass density, it must be such that the total density, $\xi_0(\mathbf{r}) + \xi'(\mathbf{r})$, is nowhere negative. Not all observable functionals vary over the surface. Mass and moment of inertia are single integral properties of the whole Earth. So also are the frequencies and amplitudes of the free elastic oscillations. Both types of problem may be written in matrix form. If the integral in the integral expression is replaced by a finite sum, η and ξ become vectors, with n and $n+m$ elements respectively, and

$$\eta = \mathbf{K} \cdot \xi,$$

where the kernel matrix, \mathbf{K} , has n rows and $n+m$ columns.

Similarly η might be the set of periods and amplitudes of free oscillations, ξ a set of values of density and elastic moduli over finite intervals within the Earth.

Evidently \mathbf{K} has no unique inverse, although the problem could be arbitrarily restricted by dividing the interior space into a number of elements not exceeding the number of observables, and estimating the mean values of ξ within each subdivision. Generally, however

that would be too restrictive. If the interior space is divided more finely, the values of the function ξ can no longer be found uniquely, but some idea of the range of possible solutions and of the reliability of any result might be found by comparing different solutions that satisfy the equations. That essentially is the course that has been followed in many applications of inverse theory.

It is usually possible to obtain one solution from a generalised inverse. Let \mathbf{K}^T be the transpose of \mathbf{K} .

Then

$$\mathbf{K}^T \cdot \eta = \mathbf{K}^T \cdot \mathbf{K} \cdot \xi.$$

Since $\mathbf{K}^T \cdot \mathbf{K}$ is square with $n+m$ rows and columns it has an inverse $(\mathbf{K}^T \cdot \mathbf{K})^{-1}$, and so one solution for ξ is

$$\xi = (\mathbf{K}^T \cdot \mathbf{K})^{-1} \mathbf{K}^T \cdot \eta.$$

$(\mathbf{K}^T \cdot \mathbf{K})^{-1} \mathbf{K}^T$ is the generalised inverse of \mathbf{K} , but the solution is not unique. ξ may be augmented by any vector ξ' which satisfies $\mathbf{K} \cdot \xi' = 0$. That is a set of n equations for $n+m$ variables to which there is no unique solution; any m variables may be chosen arbitrarily, and the remaining n are then functions of them.

There are various formal considerations that can be applied to constrain the solutions, for example, it might be that the variable ξ should be adjusted to vary as smoothly as possible within its range, conditions being applied not only to the variable itself but also to its derivatives. The earlier Jeffreys-Bullen models of the interior of the Earth that were derived from times of travel of seismic waves, were subject to conditions of that sort. Tarantola (1987) has discussed statistical principles for choosing solutions, while Parker (1994) has set out minimisation procedures for doing so. However, such constraints do not resolve the essential logical problem, for either they are based on physical considerations of how the interior of the Earth ought to be constituted, or they are adopted for purely numerical convenience, or they reflect metaphysical principles or other rules of argument.

3. Physical restrictions in inverse problems

Physical considerations are widely applied to limit the formal possibilities of solutions to inverse problems. That might at first seem ambitious, since the essence of the inverse problem is that we can only observe on boundaries and therefore do not have direct knowledge of the physical constitution of matter in the interiors of boundaries. In some instances, however, our general knowledge of the properties of materials can go far to resolve ambiguities; the most clear-cut instance is the constitution of Jupiter. We know two facts about the density distribution within Jupiter, the total mass of Jupiter and hence the average density, and the moment of inertia and hence a general indication of the increase of density from the outside in. Both are found from analyses of the motions of the natural satellites together with the spin angular velocity of Jupiter. The average density rules out any species except hydrogen as the major constituent.

The increase of density into the centre might arise from an increase in the proportion of heavy elements, but the constraint set by the average density makes it difficult to have a sufficient quantity of heavy elements to give the observed low value of the moment of inertia. It has in fact been realised for many years that the compression of hydrogen under the self-gravitation of the planet would make atomic hydrogen metallic, that is to say, it would consist of protons in a sea of unattached electrons. It has also been known for many years that the quantum mechanical theory of that simplest of all metals is elementary. The upshot is that, a few minor modifications apart, the overall properties of Jupiter can be matched by those of a planet composed almost entirely of metallic hydrogen. Although the metallic state of atomic hydrogen has not been attained in laboratory experiments, the essential conclusion is inescapable, for only hydrogen can be the main constituent of Jupiter. The only significant uncertainty is the pressure (and therefore the radius in Jupiter) at which hydrogen transforms from the molecular to the metallic form (Cook, 1980).

Jupiter also displays the limitations of theory. It has a strong magnetic field, and according to current theoretical ideas, that field would be generated by electrical currents in an electrically conducting liquid. But would metallic hydrogen be liquid in the conditions in the interior of Jupiter? It seems that is a difficult question to answer because the difference of energy between the liquid and solid forms is very small.

Jupiter is a special case because of its low density and because the rapid rotation leads to a large polar flattening which allows the moment of inertia to be rather well determined from the external value of gravity. That contrasts strongly with most other planets, apart from the Earth, for which no such clear-cut estimates of the interior states can be obtained, in part because the necessary dynamical information is lacking. Because Venus and Mercury rotate very slowly, the polar flattenings are small and therefore the moments of inertia cannot be found from gravitational observations.

When the estimates of the mass and moments of inertia of the Earth were combined with seismological studies it was realised that the Earth had a central core of much greater density than the surrounding mantle, and it was soon suggested it might consist of iron, possibly mixed with some lighter elements. In recent years shock experiments on iron have made that conclusion very probable, but they cannot rule out other possibilities. Similarly investigations of the phase transitions in olivines and pyroxenes have shown that it is very likely that there are two major divisions in the mantle, an outer one with low pressure forms of the silicates, and an inner one with the high pressure forms. Again, the experiments, taken with geochemical arguments, give strong support to such models, but cannot rule out all other possibilities.

The foregoing discussion has been about the global properties of the planets for which the mathematical form of the inverse problem is that of estimating a function within a spherical boundary from data on the boundary. There is also a local form of the inverse problem, in which a function throughout a semi-infinite space is to be estimated from data on the plane

boundary (the local surface of the Earth). That is the form of the problem encountered in exploration geophysics. One particular result that often applies is that if the data are expressed as Fourier components of functions of position on the surface, there are limits to the depths below which anything can be known about the sources of given wavelengths. In general the effect of some function of wavelength λ at a depth h below the surface dies away as $\exp[-\lambda f(h)]$, where $f(h)$ is a function of h that depends on the kernel.

A third form of the inverse problem has recently come to the fore, to which Michele Caputo has contributed. Just as the interior of the Earth was initially investigated by means of the times of travel of elastic waves through it, so the velocity structures of the oceans and atmosphere are now being investigated by the times of travel of acoustic and radio waves through them in a form of tomography (Antonelli and Caputo, 1994). Tomography is also well known in its medical applications using acoustic waves and X-rays. The observed times of travel are functionals of the velocity distributions in the sea or atmosphere or human body, and the interpretation of them is again an inverse problem. While the aim of medical studies is to obtain information about the inside of a patient, and that of the marine studies is to derive such properties as the temperature and salinity structures of the oceans and their possible changes with time, one aim of the atmospheric studies is somewhat different, it is to enable times of travel of electromagnetic signals from spacecraft to be corrected for the atmospheric conditions. Since any solution that does not affect the travel times is now of no interest, the conditions on the inversion are much relaxed.

5. Functionals from functionals

The allowable solutions to an inverse problem have been most thoroughly considered for the interpretation of the free oscillations of the Earth, not surprisingly since it was there that the essential features of the inverse problem were first explicitly encountered and analysed

(for example, Backus and Gilbert, 1968). An important concept developed in those studies was that of the trade-off between complementary properties of the distributions of densities and elastic properties that were sought, an analogue of Heisenberg's uncertainty principle. Supposing the radius of the Earth divided into a number of ranges, then the greater the number of ranges, the greater the uncertainty of the estimates of properties within any range. Conversely, if the properties were estimated in just a few ranges, then the uncertainties of the mean values within each range were less.

Just two mechanical properties, the mass and the principal moments of inertia, can be derived from dynamical observations of a planet. Somewhat more can be obtained from the variation of the attraction of gravity over the surface. Those examples, together with the analysis of the treatment of free oscillations, suggest that it is not profitable to overcome the inherent obstacles to finding functions that will give the observed functionals, but rather to look for properties of those functionals, functionals of functionals, that might be related, as the mass and moments of inertia can sometimes be, to definite physical or material properties. Put another way, combinations of physical properties should be sought that can be related to definite functionals of functionals. Tomographic studies of the atmosphere afford an example: the aim is not so much to find the distribution of the velocities of radio waves in the atmosphere, as to be able to calculate the time of travel from an arbitrary spacecraft.

6. Chaotic dynamics

At the end of the nineteenth century and the beginning of the twentieth, Newtonian dynamics was extended in three ways, each of which revealed limitations on possible knowledge of the physical world. Relativity showed that it was not possible to determine various properties, such as time and distance, or wavelength and frequency, separately for distant events, but only linear combinations of them related by the Lorentz transformation. That was the consequence of the fact that information about

distant events can only be obtained from electromagnetic signals. Quantum mechanics also set limits to what quantities could be observed simultaneously. The theoretical structures of both relativity and quantum mechanics follow from the limitations on the observations that we can make, and from the ways in which we define standards, such as that of time, to which our observations are referred. The impact of the third development, of chaotic dynamics, on physics has been appreciated less immediately. Although Poincaré, who first developed the principles of chaotic dynamics in his studies of celestial mechanics (Poincaré, 1928), stated the position quite clearly:

«A very small change that escapes our notice determines a considerable effect that we cannot fail to see, and then we say that that effect is due to chance. If we knew exactly the laws of nature and the situation of the universe at the initial moment, we could predict exactly the situation of that same universe at a succeeding moment. But even if it were the case that the natural laws no longer had any secret from us, we could still only know the initial situation approximately. If that enabled us to predict the succeeding situation with the same approximation, that is all we require, and we should say that the phenomenon had been predicted, that it is governed by laws. But it is not always so; it may happen that very small changes in the initial conditions produce very great ones in the final phenomena. A small error in the former will produce an enormous error in the latter. Prediction becomes impossible and we have the fortuitous phenomenon.»

Again, in an article on meteorology:

«Why have meteorologists such difficulty in predicting the weather with any certainty...? We see that great disturbances are generally produced in regions where the atmosphere is in unstable equilibrium. The meteorologists see very well that the equilibrium is unstable, that a cyclone will be formed somewhere, but exactly where they are not in a position to say; a tenth of a degree more or less at any given point and the cyclone will burst here and not there, and extend its ravages over distances it would otherwise have spared... Here, again, we find the same contrast between a trifling cause which is inappreciable to the observer, and considerable effects, that are sometimes terrible disasters.» (see Poincaré, 1905)

Chaotic dynamics now has a very large literature (for example, Drazin, 1992; Medio, 1992; Ott, 1993) and the circumstances that lead to chaotic behaviour are rather well understood. If the differential equations that determine the evolution of a system in time are linear, then there is a commensurable relation, as Poincaré expressed it, between the initial and final states. If the equations are non-linear, the relation may be incommensurable and may change as the system evolves in time. That is probably the most generally recognised condition. Another condition is that there may be bounds on the variables, for when a variable attains a bounding value, the memory of earlier initial conditions may be lost when it deviates from the boundary. Lastly, a large number of effective variables may lead to chaos, for as the number of variables increases, so in general does the number of ways in which a system may evolve and therefore the greater the precision with which the initial conditions must be stated for the behaviour to be predictable. With no more than six or seven variables, the necessary precision is already unrealistic.

Chaotic behaviour is not necessarily entirely random. There may be fixed points where the rates of change all vanish. When a system attains such a state, which is not the same as a state of equilibrium, a minimum of potential in linear dynamics, any memory of the initial conditions will be lost, since the state may be reached from a range of initial states. There may be so-called basins of attraction, about which a system oscillates in an almost periodic fashion. The statistical properties of chaotic time series are not the same as those of purely random phenomena, and certain statistical properties may be determined by the properties of the governing equations or because of constraints such as conservation of energy or angular momentum. Nonetheless, the most striking manifestations of chaotic behaviour are those that appear unpredictable. There are many examples of systems where quite small changes in the parameters of the system or in external conditions may lead to unpredictable shifts between alternative stationary states. Another important type of behaviour is the transition from regular to wholly random motion.

That is particularly so in hydrodynamics, the physics of which is essentially non-linear on account of the Coriolis terms. Motion which begins smoothly develops periodic oscillations, at first with just one or two periods but then with more and more until the motion is fully turbulent.

Besides the hydrodynamical non-linearities in the behaviour of the atmosphere, there are others in the transport of heat and moisture and in the chemical reactions in the atmosphere. Thus the atmosphere behaves chaotically and it is now known that the weather is inherently unpredictable beyond about three weeks; the longer term behaviour, that we call climate, is predictable over much greater times. One of the reasons for that is that climate is controlled to a considerable extent by the state of the oceans which transport heat, water, carbon dioxide and plants and animals much more effectively than the atmosphere, but much more slowly. The oceans also behave chaotically, and Poincaré's remarks on the origins of cyclones in the atmosphere apply equally to the generation of great gyres in the oceans, but on a longer time scale.

The origin of the Earth's main magnetic field (like those in other celestial objects) most probably lies in dynamo action in a liquid electrically conducting core. The equations that control it are inherently non-linear, both hydrodynamic and with interactions between magnetic fields and fluid flows and electrical currents and magnetic fields. The spin of the Earth ensures that the exterior dipole field, as observed at a distance from the dynamo source, is aligned nearly but not exactly with the polar axis, but both senses are possible. As is well known, the field switches erratically from north to south, while maintaining approximately the same magnitude. That is entirely characteristic of a chaotic system.

More than forty years ago Rikitake(1958) devised a very simple mathematical model of a self-sustaining dynamo, which gave erratic reversals of the magnetic field, with attractive similarities to the behaviour of the Earth's field. It was evidently far too simple a system to be a true representation of the Earth's dynamo. The similarity between the behaviour of

Rikitake's model and the Earth raises an issue which has also been brought up by attempts to model various biological processes by chaotic systems that are much simpler than real populations or organisms. It appears that the essential features can be represented by mathematical structures with rather few variables, which therefore are much more general than any more realistic but complex model. Consequently the observed behaviour reflects general properties of the structure of a system and is insensitive to the details of any actual one. The general mathematical and physical or biological characteristics of a system that allow it to be represented by a simple model have not been examined in any detail, but it is clear that specific knowledge of a particular system cannot be inferred from overall chaotic behaviour.

Somewhat related problems are raised by the representation of distributions in space or time by fractal formulae. Thus the distribution of numbers of lunar craters by size follows a fractal law, and so does the distribution of earthquakes by magnitude (see Turcotte, 1992). The reason for the distribution can be shown to lie in the physics of cratering or of seismic rupture, but while the physics is consistent with the observed distribution, the form of the distribution is so general that the physics cannot be inferred from the distribution – indeed it would probably be possible to argue from the distributions that earthquakes were caused by meteorite impacts and lunar craters by moonquakes.

7. Geophysical prediction

An essence of chaotic dynamics is that the initial conditions may have little or no influence on the subsequent behaviour of a mechanical system. Prediction, as Poincaré said, becomes impossible. The fundamental point is that the initial values of the state variables and those of a later state are incommensurable, that there is no unique relation between them. In prediction we know the present state but later states cannot be calculated unambiguously from the controlling equations.

Forecasting the weather and predicting climate are important activities restricted by the chaotic behaviour of the atmosphere and oceans. Chaotic dynamics may also illuminate the prediction of earthquakes. A purely statistical approach seems unsatisfactory. Small earthquakes occur in sufficiently large numbers that the probabilities of their occurrences can be mapped, but large earthquakes are so infrequent that no meaningful statistics can be established. Apart from the paucity of events, the large earthquake may so change the state of stress and the geometry of the hypocentric region that successive large earthquakes do not all come from a single population with unchanging characteristics.

A large earthquake has features that suggest that it may be the result of a chaotic process, in that it is an abrupt change from one more or less stationary state to another. The simplest model is of plate movement developing stresses that eventually exceed the strength of the crust of the Earth, leading to rupture (see Turcotte 1992). Various features of earthquakes indicate that that is too simple a model. The stresses can be and are relieved in many instances by sequences of small earthquakes, how is that stresses ever build up sufficiently to generate a large earthquake? So soon as complications are added to the simplest model, the conditions for chaotic dynamics begin to appear.

Sudden changes of state of a chaotic system do not take place when some specific condition, such as a limiting stress, are reached, for then they might be predictable. Some small insignificant event brings them about and so they are unreplicable.

8. Geophysical history

Historical geophysics goes back at least to the seventeenth century when Edmond Halley suggested, from historical records of eclipses, that the Moon was speeding up in its orbit, and when he estimated an age of the Earth from the saltiness of the sea and the rate at which rivers brought in salt (Halley, 1693, 1695, 1715). Today we seek to reconstruct the forms of conti-

nents in past times from palaeomagnetic observations, or past oceanic circulations and temperatures from compositions of deposits they laid down. The origin of the solar system remains after many centuries and after a mass of new information from space programmes, still mysterious. So too do the origin of the Moon, the early history of the Earth and how continents and oceans came into being. New information that might be expected to solve such questions often opens up fresh ones. It is hardly surprising, for although the Earth preserves considerable fragments of records of its physical and biological history, they, like written documents in archives, are not the whole story, they may have been corrupted, and they have to be interpreted in the light of whatever may be known of how they came into being.

We, like Halley, hope to get back to the past by inferences from the present state according to known physical principles, but no conclusion can be certain. Many physical processes that appear to be deterministic may also be to some slight extent chaotic, even if that is undetectable by our experiments and observations. At the same time, some of the processes to which appeal is made are sufficiently complex to be chaotic. Consequently, inferences made about the past from present conditions almost certainly involve chaotic dynamics. It is even more certain that the dynamics of biological processes are chaotic.

If there is no unambiguous relation between initial conditions and subsequent behaviour, then not only is prediction not possible, but also a former state cannot be inferred from a latter. Chaotic dynamics therefore limits historical geophysics, what we may learn of past conditions from the present. The most we are justified in saying is that some past state of affairs is possible, or probable, not that it was certain.

9. Conclusions

Among the ways of resolving uncertainties in the conclusions to be drawn from geophysical data, whether or not they be recognised as such, there are sometimes appeals to currently

received ideas on how the Earth or the solar system or the cosmos is constituted, or to more metaphysical arguments about scientific method. The sciences however try to understand the real world in an empirical manner, and the true test of a scientific hypothesis is that set out long since by Robert Boyle:

«That it enable a skilfull Naturalist to foretell future Phenomena by their Congruity or Incongruity to it; and especially the Events of such Experiments as are aptly devised to Examine it; as Things that ought or ought not to be Consequent to it.» (quoted by Westfall, 1971)

The real world as we know it empirically is the only touchstone for the scientist. The ways in which theory in the physical sciences depend on the methods of observation that the natural world allows us to make are becoming increasingly clear, for example, in relativity. In a similar way the observations that it is possible to make of the Earth and the planets have their characteristic limitations which should determine the attainable aims of geophysics and the theoretical forms in which they are expressed.

REFERENCES

- ANTONELLI, P. and M. CAPUTO (1994): Tomography of the atmosphere from times of flight to satellites: methods and feasibility, in *Atti XIII Convegno GNGTS, Roma*, November 1994.
- BACKUS, G.E. and J.F. GILBERT (1968): The resolving power of gross Earth data, *Geophys. J. R. Astron. Soc.*, **16**, 169-205.
- COOK, A.H. (1980): *Interior of the Planets*, Cambridge, XII, pp. 348, Chapter 7 and 8.
- DRAZIN, P.G. (1992): *Non-Linear Dynamics*, Cambridge, XIV, pp. 317.
- HALLEY, E. (1693): Emendationes ac Notae in vetustas Albatenii Observationes Astronomicas cum restitutione Tabularum Lunisolarium ejusdem Authoris, *Philos. Trans. R. Soc.*, **17** (204), 913-921.
- HALLEY, E. (1695): Some account of the ancient state of the City of Palmyra with short remarks upon the inscriptions found there, *Philos. Trans. R. Soc.*, **19** (218), 160-175.
- HALLEY, E. (1715): A short account of the cause of the saltness of the oceans, and of the several lakes that emit no rivers; with a proposal by help thereof, to discover the Age of the World. Produced before the Royal Society by E. HALLEY, R.S., *Secr. Philos. Trans. R. Soc.*, **29** (344), 296-300.
- MEDIO, A. (1992): *Chaotic Dynamics, Theory and Application to Economics*, Cambridge, XVIII, pp. 344.
- OTT, E. (1993): *Chaos in Dynamical Systems*, Cambridge, XII, pp. 385.
- PARKER, R.L. (1994): *Geophysical Inverse Theory*, Princeton.
- POINCARÉ, J.H. (1905): *La Valeur de la Science*, Paris.
- POINCARÉ, J.H. (1928): *Sur les Courbes Définies par les Equations Différentielles*, Oeuvres, tome 1, Paris, pp. 40-222.
- RIKITAKE, T. (1958): Oscillations of a system of disc dynamos, *Proc. Camb. Philos. Soc.*, **54**, 89-105.
- TARANTOLA, A. (1987): *Inverse Problem Theory: Methods for Fitting and Model Parameter Estimation*, New York.
- TURCOTTE, D.L. (1992): *Fractals and Chaos in Geology and Geophysics*, Cambridge, pp. 221.
- WESTFALL, R.S. (1971): *The Construction of Modern Science* (Wiley and Sons.), XIV, pp. 171; p. 115.