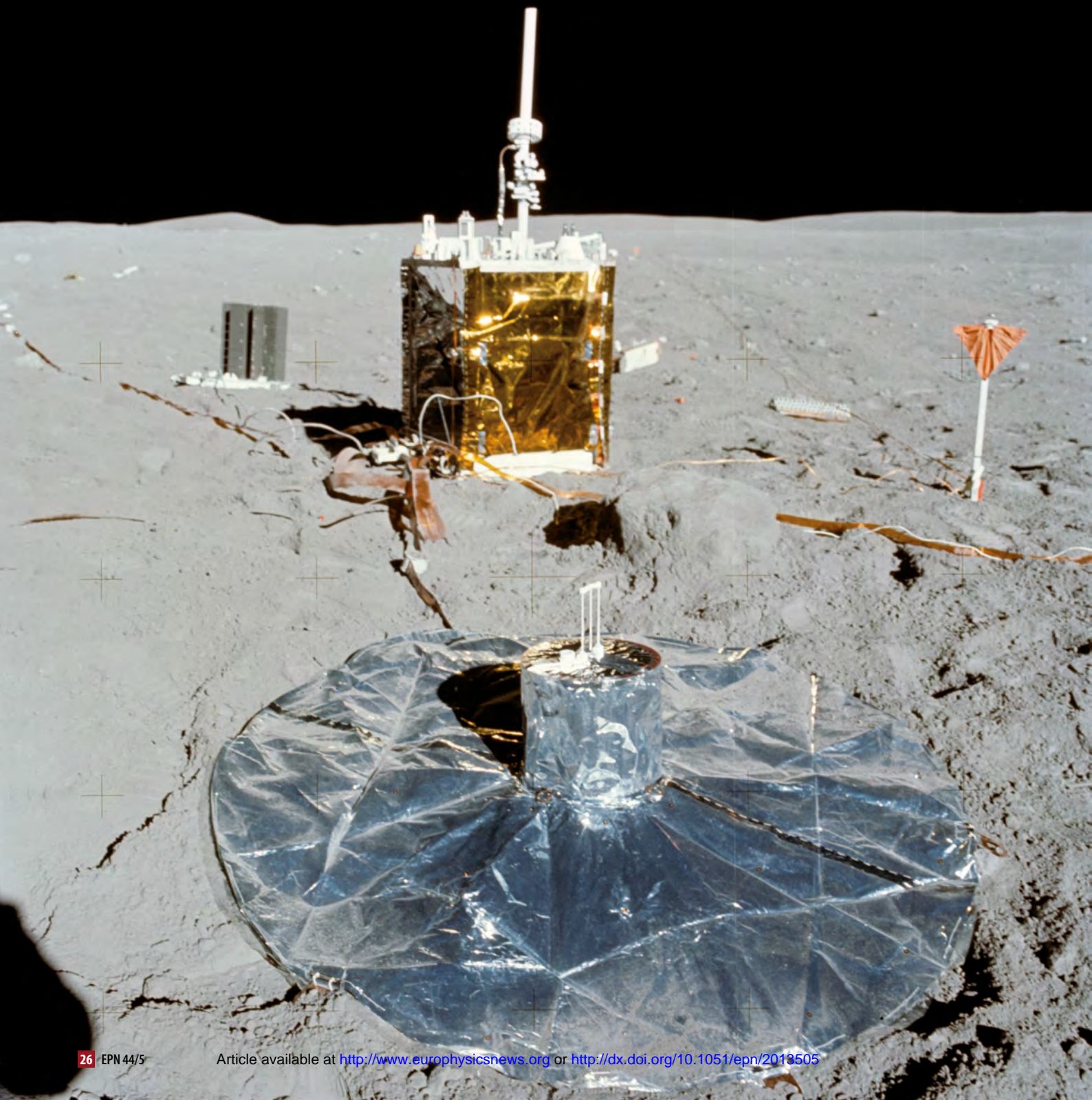


ON INVERSE PROBLEMS IN PHYSICS

■ Georges Jobert – email: jobert.g@wanadoo.fr – DOI: [10.1051/epr/2013505](https://doi.org/10.1051/epr/2013505)



Despite its great practical importance, the theory of inverse problems remains poorly known. Indeed, physics was built by solving *direct* problems: from a given model to the corresponding measurable data. The opposite approach: from observed data to model parameters – the solution of an ‘inverse’ problem – is not so easy. It is, however, essential in many fields.

The announcement made on September 22, 2011 by researchers of the OPERA [1] experiment, that the speed of observed neutrinos seemed to exceed the speed of light, acted as a bombshell, not only in scientific circles, but also in the press. Our purpose is to show that a very general theory was discreetly at work here, as in dozens of applications in everyday life. Let us first recall that the experiment was based on the creation of neutrinos at CERN near Geneva and on the detection of their arrival at a station under the Gran Sasso in Abruzzo. Neutrinos followed straight-line paths in the terrestrial crust, the length of which – about 730 km – was determined by GPS measurements having extreme precision [2]. But GPS tracking, medical engineering (echography, tomography,...), detection of flaws in devices by non-destructive techniques, pattern and characters recognition, detection of new planets, determination of the internal structure of the Earth [3] and stars, oil and gas research, and many other applications – from the diffraction of particles in quantum mechanics [4] to econometrics – would not be possible without the progress made during the past fifty years in the theory of inverse problems (and without the explosion of computing power required for their solution). A striking example of the determination of the internal structure of Tibet is shown in Fig. 1. In the case of GPS, it has been first necessary to solve the inverse problem of determining the gravity field of the Earth from the observation of trajectories of many satellites (first the Moon, then artificial satellites). Once this field known, the trajectories of a group of satellites equipped with transmitters and very accurate clocks can be determined. The position of a point on the surface of the Earth can then be calculated from the time of arrival of the waves transmitted by the satellites. Actually, improving the models of the

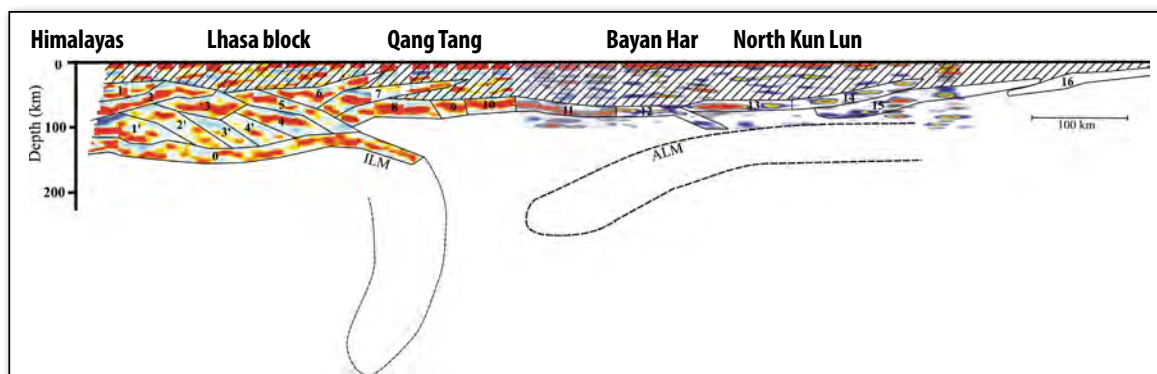
gravity field is obtained by simultaneously treating both problems: computing trajectories and positioning stations. This is a particularly interesting aspect of the theory of Albert Tarantola, summarized in box 1 (see also ref. [9]).

On the theory of inverse problems

The usual approach of experimental science can be described as follows: The observed phenomena lead to formulate hypotheses and to develop a theory, in which a model of the system is built, characterized by some relevant parameters. The theory is then used to solve the ‘direct’ problem, *i.e.*, to calculate, from the values chosen for these parameters, the values of the measurable data, which will be compared to experiment. In short, a theory is the formulation of a direct problem. Consider for example a problem studied by Abel in 1826: on the line of greatest slope of a hill (assumed of known shape and perfectly smooth) a vehicle is launched upwards. Knowing the profile and the initial velocity of the vehicle, find the duration of a round trip. By comparing the calculated values to the observed ones (or to the result of experiments programmed for this purpose) one can check the validity of the theory.

To determine the properties of the studied system from the observed data, *i.e.*, to solve an inverse problem, is an otherwise difficult operation because it proceeds opposite to the normal process of experimental science. For the above case studied by Abel, the inverse problem is to find the shape of the hill profile, knowing the time taken for a round trip of the mobile, as a function of its initial speed. In mathematics it is, for example, the famous problem of the drum: find its form knowing the sound that it produces. One has then to search the coefficients of a differential equation or of a PDE, knowing the spectrum of the eigenvalues. One encounters this problem in particular when one wants

◀ Seismometer on the Moon (©NASA)



◀ FIG. 1: S-N section of Tibet obtained from seismological INSU expeditions between 1981 and 1998 (A.Hirn). The base of the crust (Mohorovicic's discontinuity, between 50 and 70 km) is not plane and seems often multiple. Also deeper interfaces appear.

to determine the properties of a medium that is inaccessible to direct observation, from observations made outside the medium. This is obviously the case in internal geophysics. It is therefore not surprising that many geophysicists have contributed to the progress of this discipline, by developing a variety of methods to solve their problems.

Inverse problems in geophysics

The first inverse problem studied in this field was the determination of the source (position, time) of seismic waves, from their arrival time at a number of locations. If it can be assumed that the propagation speed is constant on a flat surface, this problem can be treated by the least squares method developed by Gauss in the early nineteenth century. Once the source is determined, the time it takes for a wave to reach a given point on the surface can be calculated. On the (almost) spherical Earth, this time depends (almost) only on the distance between the source and the station, but the apparent velocity of the signal is greater than if the propagation were along the surface. Hence another inverse problem: how to determine the wave velocity as a function of depth (if it may be assumed to be a function of depth only). In 1910, Batemann, Herglotz and Wiechert applied the method used by Abel for the equivalent problem mentioned above. In practice, however, many difficulties are encountered. The solution actually requires the calculation of the derivative of the transit time over the distance, an operation unreliable because of the large experimental uncertainties. In addition, as shown by Gerver and Markushevitch [5] in 1967, it is unique only if the velocity law meets certain conditions. But this method has shown that, overall, the velocity increases with depth, with

some discontinuities about twenty miles deep, as highlighted by Mohorovičić in 1909. Gutenberg showed the existence of the Earth's core in 1913. More detailed models were then obtained. In 1967, G. Backus and F. Gilbert developed [6] a general theory which led to a revolution in the field.

Discretisation and successive approximations

The properties of a medium depend in principle on position and time. So its full description generally requires an infinite amount of information. In practice, by studying only static problems and fragmenting the medium, discrete and finite sequences of parameters remain. For example, by defining an Earth model by the density of homogeneous elementary volumes, the attraction it exerts on its exterior may be deduced from the Newtonian theory of gravitation. One then proceeds by successive adjustments of the parameters until an agreement between calculated and observed values is reached that is consistent with the instrumental precision. This is the classic approach by successive approximations, applied until the middle of last century.

Existence, uniqueness, stability of solutions for a linear problem

Linear problems – where the data depend linearly on the parameters – are the only ones for which the theory can be fully developed. It is easy [7] to show that three main difficulties are encountered. The solution may or may not exist, may be unique or not, may be stable or not when the data or the theory are slightly varied. In the last case, the problem is called "ill-posed".

BOX 1: THE PROBABILISTIC APPROACH OF ALBERT TARANTOLA

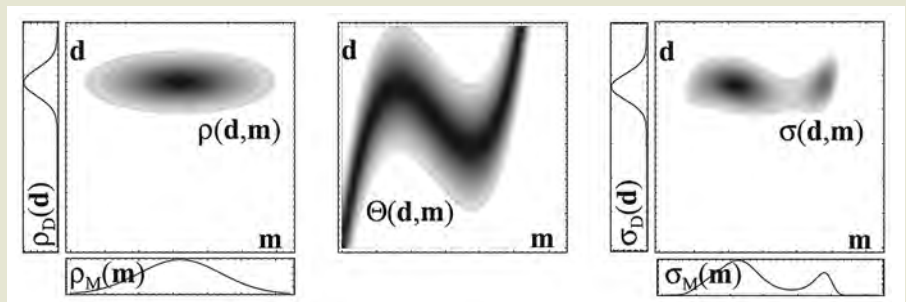
Let us assume that, before applying the inversion, we have at our disposal a probability distribution $\rho_D(\mathbf{d})$ on the observable \mathbf{d} , and some *a priori* information on the model \mathbf{m} (e.g., parameters being between certain values), represented by a law of probability $\rho_M(\mathbf{m})$. Data and parameters being independent, the *a priori* information we have on the system is represented by the product law: $\rho(\mathbf{d}, \mathbf{m}) = \rho_D(\mathbf{d}) \cdot \rho_M(\mathbf{m})$.

For the direct problem, a theory gives \mathbf{d} knowing \mathbf{m} . It may also be known only approximately – because of the physical constants used (always determined experimentally), or due to modelling errors (e.g., spherical model for the flattened Earth) – and must therefore be represented by a probability

distribution $\Theta(\mathbf{d}, \mathbf{m})$. By combining these two states of independent information one obtains the resulting probability on the data: $\rho_D(\mathbf{d}) = \Theta(\mathbf{d}, \mathbf{m}) \cdot \rho_M(\mathbf{m})$. The observation leads to the definition of the law $\sigma_D(\mathbf{d})$. The solution of the inverse problem is given by

$\sigma_M(\mathbf{m}) = \Theta(\mathbf{d}, \mathbf{m}) \cdot \sigma_D(\mathbf{d})$. We see (fig.2) how the knowledge of the model has been improved thanks to the data used. The *a posteriori* information on the system is now represented by the product law: $\sigma(\mathbf{d}, \mathbf{m}) = \sigma_D(\mathbf{d}) \cdot \sigma_M(\mathbf{m})$. Two types of solutions appear possible. ■

▼ FIG. 2: (Fig.1.12 of Tarantola's book [9]). The grey areas correspond to the values of the functions, taking the uncertainties into account.



The progress

In any case one should not look for a solution dependent on a number of parameters greater than that of the data. This is a form of the principle of parsimony (or Ockham's razor). But all the data do not have the same importance. One must take into account the information available on their probable errors, giving more weight to those that are more accurate. To this end, data covariance and parameters covariance matrices are introduced. From these a generalized reverse is built that is, in a sense, the best solution available. G. Backus and F. Gilbert have shown in particular that, before any effective measure, the number of independent pieces of information contained in a collection of data may be estimated. In principle one can thus determine, before any experiment, which ones provide the most interesting information. In parallel one can assess the resolving power which can then be obtained for the values of model parameters. A local value for a parameter can be regarded as an average weighted by a filter, the resolving power corresponding to the width of the optimal filter, which can be evaluated from the data.

The results of experimental data have an imprecision both due to errors in the data and to a finite resolving power. One can try to minimize a quantity that combines these effects. For linear problems one shows that, when the experimental error is decreased by increasing the measurement accuracy, the resolving power decreases. Inversely, searching a better localisation for a parameter leads to an increase in its error. This is a form of the uncertainty principle.

Non-linear problems

Nonlinear problems do not allow this kind of analysis. For a given model one can evaluate the difference between observed and calculated data – for example by calculating the sum of the squares (the method of least squares), or of the modules of the differences – and look for the solution in the direction where the difference decreases fastest (gradient method). However, this often leads to secondary minima which do not give the lowest value. Two other methods can be used:

a) Linearisation. If previous studies provide a result valid in first approximation, one can search a solution in the vicinity of this model. Writing the relationship between data \mathbf{D} and model \mathbf{M} as $\mathbf{D} = \mathbf{G}(\mathbf{M})$, if \mathbf{M}_0 is the initial model and ε the deviation from the final solution, we may write: $\mathbf{D} = \mathbf{G}(\mathbf{M}_0 + \varepsilon) = \mathbf{G}(\mathbf{M}_0) + \mathbf{G}'(\mathbf{M}_0)\varepsilon + \mathcal{O}(\|\varepsilon\|^2)$ where \mathbf{G}' is a function that can be constructed from \mathbf{G} . We then seek the solution of the linear problem: $\mathbf{D} = \mathbf{G}(\mathbf{M}_0) + \mathbf{G}'(\mathbf{M}_0)\varepsilon$. This process can be iterated.

b) Monte Carlo method. If the computational cost of the solution of the direct problem is not excessive, one may conduct a random exploration of the parameter domain. This avoids the difficulties mentioned above. But when the number of parameters is large, the research quickly becomes very costly.

A detailed description of these methods can be found in the works of A. Tarantola [8,9]. In ref. [9], he proposed a probabilistic approach based on Bayes (1702-1761) assumption (see box 1). In this approach, the direct problem and the inverse problem lose their specificity and are in fact replaced by the search for the state of a system. This also corresponds to the evolution of research, for example in seismology where one is no longer trying to separately determine the parameters of the seismic source (origin and focal mechanism) and those of the Earth interior. All these parameters are searched simultaneously using not only the arrival time of the waves but also the shape of the signals. For A. Tarantola, the *a priori* information on the parameters of the studied model is defined as a probability distribution on the space of models. This distribution is transformed into an *a posteriori* information, by incorporating on one hand a physical theory linking the parameters to observable quantities – which themselves should be defined by a probability distribution – and on the other, the information provided by the observed data. The knowledge of the system state is thus revised. ■

About the author



Georges Jobert, École Normale Supérieure, aggregation of mathematics, PhD (Geophysics). He worked at the *Institut de Physique du Globe de Paris*, which he directed from 1971 to 1976. He was professor at *Université Pierre et Marie Curie* (Paris) until 1983. Scientific director of CNRS (1966-1971). Corresponding member of Academy of Sciences.

References

- [1] OPERA collaboration (November 17, 2011). *Measurement of the neutrino velocity with the OPERA detector in the CNGS beam*. arXiv:1109.4897. Bibcode:2011arXiv1109.4897T.
- [2] A. Tarantola *et al.*, Relativistic corrections are required. See: "Gravimetry, Relativity, and the Global Navigation Satellite Systems", ArXiv: 0905.3798 (May 2009).
- [3] See for example fig.1 in "Des problèmes inverses en Physique", *Reflets de la Physique* **31**, 15 (2012).
- [4] K. Chadan and P.C. Sabatier, "Inverse Problems in Quantum Scattering Theory", Springer-Verlag (1977).
- [5] M.L. Gerver & V.M. Markushevitch, *Вычислительная сейсмология* **В3**, 3 (1967).
- [6] G. Backus and G. Freeman, *Geophys. J. Roy. Astr. Soc.* **13**, 247 (1967).
- [7] See for example: box 1, in "Des problèmes inverses en Physique", *Reflets de la Physique* **31**, 15 (2012).
- [8] A. Tarantola and B. Valette, *J. Geophys.* **50**, 159 (1982)
- [9] A. Tarantola, "Inverse problem theory", Elsevier, 1987, SIAM 2005. Tarantola insisted on the need to make the results insensitive to the choice of the parameterization: velocity or slowness, mass or specific volume... An idea that he developed in his last book: "Elements of Physics, Quantities, Qualities and Intrinsic Theories", Springer (2006).