Comptes Rendus (Doklady) de l'Académie des Sciences de l'URSS
1943. Volume XXXIX, № 5

MATHEMATICAL PHYSICS

ON THE STABILITY OF INVERSE PROBLEMS

By A. N. TIKHONOV, Corresponding Member of the Academy

Many objects of Nature have properties that are not well fitted to, or even altogether defy, direct investigation. In such cases some of their characteristics are studied whose manifestation can be measured. Our judgment on the structure of the earth crust, for instance, may in certain cases be based on the investigation of its characteristics, such as density, or electric conductivity, which determine the respective physical field accessible to measurement at the surface of the earth—the gravitational field, the field of electric current.

Given a certain characteristics of the medium (distribution of density or of electric conductivity), we are usually able to compute (precisely or approximately) the physical field determined by this structure in the region where the measurements are made. Yet, the problem to be solved here is the reverse to this. Namely, the physical field is known, while the structure of the medium which determines it is sought.

The customary way to solve the inverse problems is by selection. Within an arbitrarily chosen (sufficiently wide) class of possible structures of the medium the corresponding physical fields are computed and the solution of the problem arrived at by selecting some admissible medium, for which the calculated physical field shows but a small deviation from observation.

In order to put the method of selection upon a solid foundation, it is necessary to establish (or admit) the existence of certain regularities: 1) One has to establish the uniqueness theorem for the direct correspondence, i. e. to prove that no two different types of medium have a single corresponding field. Then we have also the right to speak of a reverse correspondence. Without this the method of selection has no sense at all. 2) The coincidence between the calculated and observed field is not an absolute one (if only because the selection is made in an approximate way). We therefore have, moreover, to prove the stability of the inverse problem (or the continuity of the inverse mapping), that is, to make sure that with slight deviation of the auxiliary field from observation the respective structure of the medium cannot possibly deviate strongly from the actual.

In studying the stability of the inverse problem, a number of questions of qualitative and quantitative character may be raised.

In the present paper set theoretical conceptions are applied to one of these problems, which consists in proving that under certain conditions the stability of the inverse problem is a direct consequence of the uniqueness theorem. We shall also apply it to the inverse problem of the potential and to the study of the continuous dependence of the solutions of ordinary differential equations upon the parameter.

1. In the theory of continuous mappings there takes place the following theorem [1, 2].

Let a set of elements $\{x\}$ forming a metric space R be mapped continuously upon another set of elements $\{x^*\}$ forming a metric space R^*. If the mapping $x^* = f(x)$

is one-to-one and continuous and if the space R is compact, then the inverse mapping $x = f^{-1}(x^*)$ is also continuous.

Definition. Let a set of elements $\{x\}$ be mapped by a function $f(x)$ on to another set of elements $\{x^*\}$: $x^* = f(x)$. This mapping is said to be one-to-one at the point x_0 if $x_0^* = f(x_0) \neq f(x)$ for any element x distinct from x_0.

It is easy to prove the following theorem.

Let a metric space R be mapped continuously on to another metric space R^*

$$x^* = f(x) \qquad (x \in R, \quad x^* \in R^*).$$

If this mapping is one-to-one at the point x_0 and the space R is compact, then the inverse mapping $x = f^{-1}(x^*)$ is likewise continuous at the point x_0^*.

The continuity of the inverse mapping is understood to mean that for any $\varepsilon > 0$ there exists a $\delta(\varepsilon)$ such that if

$$\rho(x^*, x_0^*) < \delta(\varepsilon),$$

then

$$\rho(x, x_0) < \varepsilon$$

where x is any prototype of the point x^*.

Though we formulate these theorems for metric spaces, they also hold in a more general case.

2. Among the direct problems of the potential theory there is one that consists in computing for the surface $z = 0$ the potential of a bounded body T filled up with a homogeneous mass of density μ and lying beneath that surface ($z < 0$). We shall demonstrate that the inverse problem has a stable solution.

Suppose approximately that the position of the disturbing body T is known to be inside a given surface S. Let us examine the totality of bodies $\{T\}$ satisfying the following conditions:

1°. Each of the bodies T belongs to a given bounded surface S lying in the region $z < 0$.

2°. Each of the bodies T is stellate with respect to its centre of gravity, so that the equation of the surface Σ, bounding the body T, may in the spherical system of coordinates with their centre at the point P be represented in the form

$$z = f(\varphi, \theta).$$

3°. Function $f(\varphi, \theta)$ has its derivative numbers bounded by a number M, common to all bodies of the class R.

Let us determine the degree of proximity of two different bodies T_1 and T_2 from the class R by means of the number

$$\rho(T_1, T_2) = \max\{\rho(P_1, P_2), \max|f_1(\varphi, \theta) - f_2(\varphi, \theta)|\}$$

where f_1 and f_2 are functions defining the surface equations of the bodies T_1 and T_2 with respect to their centres of gravity P_1 and P_2.

We shall prove the following theorem.

Whatever may be the degree of accuracy of ε and the class R of the bodies, such number $\delta(\varepsilon)$ may be indicated that if the values of the potentials (or of their derivatives) $V_1(x, y)$ and $V_2(x, y)$ for any two bodies T_1 and T_2 of the class R differ from each other at $z = 0$ by less than $\delta(\varepsilon)$

$$|V_1(x, y) - V_2(x, y)| < \delta(\varepsilon),$$

then these bodies are separated by a distance less than ε

$$\rho(T_1, T_2) < \varepsilon.$$

In fact, with the notion of distance determined as $\rho(T_1, T_2)$, the totality of bodies of the considered class R form a metric space. If the distance between the potentials (or their derivatives) $V_1(x, y)$ and $V_2(x, y)$ be evaluated by the number

$$\rho(V_1, V_2) = \max|V_1(x, y) - V_2(x, y)|$$

(the maximum exists indeed), then the set of functions $\{V\}$ presents a metric space R^*.

The correspondence between the bodies T and the values of their potentials (or derivatives) determines the mapping

$$V = f(T)$$

of the space R on to the space R^*.

This mapping is continuous and one-to-one, for, in virtue of a well known theorem by P. S. Novikov, no two different bodies T_1 and T_2, stellate with respect to their centre of gravity, may have equal potentials corresponding to them. And the centre of gravity has its position determined by the value of the potential at the surface [3].

The class under consideration, R, forms a compact family. Direct application of the theorem mentioned in § 1 just proves the theorem.

Modifying the notion of distance between the bodies or that of proximity of potentials (or of potential derivatives), one may easily establish other theorems of stability.

It is well to note that without the conditions of type 1° or type 3° the stability of the inverse problem does not hold any longer.

3. Carathéodory has established the following theorem on the continuous dependence of the solutions of a system of ordinary differential equations upon the parameter [4].

Let the functions $f_k(x; y_1, \ldots, y_n; t)$, $(k = 1, \ldots, n)$ be given, satisfying the conditions:

a) For every value of t inside a certain neighbourhood U_{t_0} of the point t_0 every function $f_k(x; y_1, \ldots, y_n; t)$ is measurable with respect to x when y_k are fixed, and is continuous with respect to $y_1, \ldots, y_n$ when x is fixed, x varying within the region $a < x < b$; and, moreover, there exists such a measurable (in the region $a < x < b$) function, independent of t, that

$$|f_k(x; y_1, \ldots, y_n; t)| < M(x) \quad (k = 1, \ldots, n).$$

b) For $i = t_0$ and arbitrary values of $y_1, \ldots, y_n$ the functions $f_k(x; y_1, \ldots, y_n; t)$ are continuous with respect to $y_1, \ldots, y_n; t$.

c) For $t = t_0$ and $\alpha_k^0(t_0) = \mathring{\alpha}_k$ there exists a sole system of functions $y_k(x, t_0) = y_k^{(0)}(x)$ satisfying the equations

$$y_k(x, t) = \alpha_k(t) + \int_x^x f_k(t; y_1(x, t), \ldots, y_n(x, t); t)\, dx \quad (k = 1, \ldots, n). \tag{1}$$

If, under these conditions,

$$\lim_{t \to t_0} \alpha_k(t) = \alpha_k^{(0)}, \tag{2}$$

then for any system of solutions (1) at any point of the region under consideration

$$\lim_{t \to t_0} y_k(x, t) = y_k^{(0)}(x).$$

Let us examine the totality of all possible systems of solutions of equation (1) when $t \subset U_{t_0}$. Defining the distance between the elements $Y_1 = \{y_1^{(1)}(x), \ldots, y_n^{(1)}(x)\}$, $Y_2 = \{y_1^{(2)}(x), \ldots, y_n^{(2)}(x)\}$ as

$$\rho(Y_1, Y_2) = \sqrt{[y_1^{(1)}(x) - y_1^{(2)}(x)]^2 + \ldots + [y_n^{(1)}(x) - y_n^{(2)}(x)]^2}$$

we obtain a metric space R which is compact by virtue of the obvious fact that the functions $y_k(x)$ are uniformly bounded and uniformly continuous in their totality [this is a consequence of the existence of $M(x)$ and the condition (2) if only U_{t_0} is small enough].

To every element $Y=\{y_1(x), \ldots, y_n(x)\}$ of the space R we make correspond a point of an n-dimensional Euclidean space $\{\alpha_1, \ldots, \alpha_n\}$, putting

$$\alpha_k = y_k(x_0) \quad (k=1, \ldots, n).$$

The distance between the images of two elements being less than that between the elements, it will be obvious that this is a continuous mapping.

In virtue of the conditions of the theorem this mapping is also one-to-one at the point $\{y_1(x, t_0), \ldots, y_n(x, t_0)\}$.

Applying the theorem of § 1, one sees that the inverse correspondence is continuous in the sense that for any ε there exists such $\delta(\varepsilon)$ that if

$$\sqrt{(\alpha_1^{(0)}-\alpha_1)^2+\ldots+(\alpha_n^{(0)}-\alpha_n)^2} < \delta(\varepsilon),$$

then, whatever the solution $\{y_1(x), \ldots, y_n(x)\}$ corresponding to $\{\alpha_1, \ldots, \alpha_n\}: y_k(x_0)=\alpha_k$ may be, we shall have

$$\sqrt{[y_1^{(0)}(x)-y_1(x)]^2+\ldots+[y_n^{(0)}(x)-y_n(x)]^2} < \varepsilon$$

and this proves the theorem.

Many examples can easily be put on the stability of inverse problems.

Received
12.IV.1943.

REFERENCES

[1] Х а у с д о р ф, Теория множеств, стр. 143 (1937). [2] A l e x a n d r o f f - H o p f, Topologie, I, p. 95 (1935). [3] P. S. N o v i k o v, C. R. Acad. Sci. URSS, XVIII, No. 3(1938); G. A. G a m b u r z e v, Bull. Acad. Sci. URSS, sér. géophys., No. 4 (1938); A. A. Z a m o r e v, ibid., No. 3 (1939). [4] C a r a t h é o d o r y, Vorlesungen über reelle Funktionen, S. 678 (1918).