

G. A. Viano: ANALYTICITY AND NOISE IN THE NUMERICAL
ANALYTIC CONTINUATION OF SCATTERING DATA.

Recently the relevance, in particle physics, of the improperly-posed problems has been clearly pointed out (see ref. (1) and the papers there quoted). More precisely, in many cases, the needed information can be extracted from the analysis of the scattering data through the numerical analytic continuation. Therefore one must take into account two opposite aspects of the problem: the analyticity of the functions and the noise of the experimental data.

Let us observe that, in many instances, to extract information through the numerical analytic continuation of the scattering data is an inverse-problem in a rather generalized sense. In the present note, following Tikhonov and ref. (2), we call direct-problems those which are oriented along a cause-effect sequence (i. e. problems of finding out the consequences of given causes); in this sense the inverse-problems are those associated with the reversal of the chain of causally related effects, i. e. problems of finding the unknown causes of known consequences. In other words, assuming the previous definition, not only the reconstruction of the potential from the knowledge of the angular distribution (non-relativistic inverse-problems at fixed energy), but also the determination of a coupling constant or of the spectral function in a dispersion relation and so on, are inverse-problems in a generalized sense. From this point of view these problems look very similar to the reconstruction of the input-signal from the reaction at the output of a device^(x). In this sense, some of S-matrix methods extensively used in particle-physics, appear to be a particular case of the more general methods of system-theory. Nor this is surprising,

(x) - More precisely, in our case, we are not interested to the reconstruction of the input but to the device.

since, as it is well-known, the dispersion relation approach was originally applied in electrical network analysis and in optics and successively used in particle-physics;

Now many inverse-problems of mathematical physics are improperly-posed since the continuity of the solutions on the data is lacking. The more appropriate way of approaching these ill-posed problems seems to be a probabilistic method, which should take into account the random nature of the noise. At this purpose Turchin et al.⁽²⁾ tried to use mathematical-statistics methods for solving the incorrectly-posed inverse-problem related to the solutions of the Fredholm integral equation of the first kind; the methods of Turchin et al. seem to us extremely useful and promising. On the other hand many authors, and particularly R. Cutkosky⁽³⁾, worked out statistical methods for the numerical analytic continuation of scattering data; however the Cutkosky procedure looks rather cumbersome and it is not very clear how it is related to the non-statistical regularizing procedures. Therefore the main purpose of the present note consists in applying the method of Turchin et al. to the specific case of the analytic continuation of complex-valued functions, holomorphic in a certain domain, and in showing the relationships which exist between this probabilistic approach and the methods which have been extensively discussed in ref. (1), which will be referred hereafter as I.

Now we must briefly recall the general formulation of the problem; in this part we shall strictly follow I. We shall consider essentially the simplest one of the many examples, which have been analyzed in I; this fact, however, is not restrictive since the extension to the other examples is straightforward. Therefore we consider the geometry of the unit disc Ω in the complex z -plane; in fact the geometries of the problems of physical interest can always be mapped, thanks to the Riemann mapping theorem, into that of the unit disc. The problem is to approximately determine by analytic continuation certain values of a function $f^0(z)$ which is holomorphic in Ω , but where measurements for f^0 are possible only at data points in the segment $\Gamma = [-a, b]$ of the real axis, which segment is called the physical region. We let $\tilde{\Gamma}$, the data set, denote the set of data points actually available and used in any particular instance. In other words we desire to approximately determine an unknown element f^0 of a certain space X of functions f holomorphic on the complex domain Ω ; we call X the solution space. For each f in X , Af denotes a certain function on the data set $\tilde{\Gamma}$, which can be approximately measured physically. Let h denote the data function actually measured on $\tilde{\Gamma}$ as our approximation to Af^0 . The possible Af and h lie in a certain space Y of functions on $\tilde{\Gamma}$; we call A the data operator and Y the data space. In order to stabilize the problem one usually requires some sort of global bound on f^0 ; in most cases this can be interpreted as a bound on the boundary value of f^0 . For each f in X , let Bf denote certain "boundary-values" of f on $\partial\Omega$. The possible Bf lie in a certain space Z of functions (or generalized functions)

on $\partial\Omega$; we call B the constraint operator and Z the constraint space.

Now we investigate how it is possible to get stability for the solutions in our case. At this purpose let us recall the following compactness theorem (see ref. (4), p. 141).

Theorem: Let σ be a continuous map on a compact topological space; if σ is 1-1 then its inverse σ^{-1} is continuous.

Let us assume that the functions f , holomorphic in Ω , satisfy the following bound

$$(1) \quad \left\| Bf \right\|_Z = \sup_{z \in \partial\Omega} |f(z)| \leq E$$

This condition implies that the functions $f(z) \in \mathcal{F}$ (\mathcal{F} is the family of functions holomorphic in Ω and satisfying condition (1)) are uniformly bounded on every compact subset of Ω . Therefore we can say, thanks to the Montel theorem (see ref. (5), p. 141), that the family \mathcal{F} is normal; i. e. from any sequence $f_1(z), f_2(z), f_3(z), \dots$ of functions of \mathcal{F} it is possible to extract a subsequence which converges uniformly in any closed subdomain of Ω . This class of functions is, moreover, compact. Indeed, if $\{f_n(z)\}$ is a uniformly converging sequence of such functions, it follows from $|f_n(z)| \leq E$ that the limit $f(z)$ of this sequence must also satisfy $|f(z)| \leq E$ at all points of Ω . But, as the limit of a uniformly converging sequence of functions which are regular in Ω , $f(z)$ is also an analytic function regular in this domain; $f(z)$ belongs therefore to the original class of functions, which has thus been shown to be compact. This fact, plus the uniqueness of the analytic continuation restores continuity to the problem of analytic continuation from the physical region to any closed subdomain of Ω . Of course the condition (1) is not sufficient if we want to perform a continuation up to the boundary of Ω . In such a case we must reinforce the bound, requiring, for example, that $\sup_{z \in \partial\Omega} \left| \frac{df}{dz} \right|$ is bounded.

In this way we can guarantee the continuity up to the boundary, even if the restored continuity is extremely weak; i. e. of logarithmic type^(1, 6).

In general we assume that X, Y, Z are linear spaces, with norms $\| \cdot \|_X, \| \cdot \|_Y, \| \cdot \|_Z$; that a is a continuous linear operator mapping X into Y , and that B is a continuous linear operator mapping X into Z . We let $\langle \cdot \rangle$ denote a norm or seminorm which will be employed to measure the solution accuracy.

Then the general problem can be formulated as follows.

Problem: Suppose that f^0 satisfies

$$(2) \quad \left\| Af^0 - h \right\|_Y \leq \varepsilon$$

4.

$$(3) \quad \left\| Bf^0 \right\|_Z \leq E$$

where E is a "fixed" number and ε is a "small" number. We assume that both ε and E are known; We want to find an element f^1 in X which approximates f^0 , in the sense that $\langle f^1 - f^0 \rangle$ is small when ε is small (and the number of data points in Γ is sufficiently high).

For the sake of simplicity we shall suppose that X , Y and Z are real Hilbert spaces (for details see I). Now if f^0 satisfies (2) and (3), then it also satisfies

$$(4) \quad \left\| Af^0 - h \right\|_Y^2 + \left(\frac{\varepsilon}{E}\right)^2 \left\| Bf^0 \right\|_Z^2 \leq 2\varepsilon^2$$

Conversely, any f satisfying (4), satisfies (2) and (3) except for a factor of at most $\sqrt{2}$. At this point we can recall the first method of I for obtaining an approximation f^1 . Let our approximation f^1 be that element of X which minimizes

$$(5) \quad \left\| Af - h \right\|_Y^2 + \lambda^2 \left\| Bf \right\|_Z^2$$

with $\lambda = \varepsilon/E$. It is the solution of the normal equation

$$(6) \quad Cf = (A^*A + \lambda^2 B^*B)f = A^*h$$

In the present note we do not discuss further the solution of eq. (6) and the properties of the operator C (we limit to recall that the normal operator C has a bounded inverse); for all these questions the reader is referred to I.

Before going to the probabilistic approach it is better to reconsider the problem in a discretized form. At this purpose we replace the infinite dimensional space X by an n -dimensional approximating subspace $X_n \subset X$ with basis elements $\varphi^1, \dots, \varphi^n$. By approximating subspace we mean that every f of interest has an approximation \underline{f} in X_n such that $A\underline{f} \approx Af$, $\langle f - \underline{f} \rangle \approx 0$, and $\left\| B\underline{f} \right\|_Z \approx \left\| Bf \right\|_Z$. However, we are going to neglect the discretization errors, since they can be made negligibly small by making n sufficiently large. That is, we assume that all f of interest actually lie in X_n . One then merely proceed with the above method but with X replaced by X_n . Then all computations reduce to linear calculations involving the basis coefficients. The general f has the form

$$(7) \quad f = x_1 \varphi^1 + x_2 \varphi^2 + \dots + x_n \varphi^n$$

where the vectors x vary over the space R^n of all n -dimensional

column vectors with real components x_j . We equip R^n with its usual Euclidean inner product and norm. Moreover we define maps \mathcal{A} from R^n into Y and \mathcal{B} from R^n into Z as follows:

$$(8) \quad \mathcal{A} x = A (x_1 \varphi^1 + x_2 \varphi^2 + \dots + x_n \varphi^n)$$

$$(9) \quad \mathcal{B} x = B (x_1 \varphi^1 + x_2 \varphi^2 + \dots + x_n \varphi^n)$$

Conditions (2) and (3) now become

$$(10) \quad \|\mathcal{A} x^0 - h\|_Y \leq \varepsilon$$

$$(11) \quad \|\mathcal{B} x^0\|_Z \leq E$$

where x^0 is the coefficient vector of f^0 . Finally the normal equation will assume the following form:

$$(12) \quad \mathcal{E} x \equiv (A^* A + \lambda^2 B^* B) x = A^* h$$

Now we can analyze the problem from the probabilistic point of view. Therefore the solution and the data vectors must be regarded as random vectors. In this note we shall use Gaussian distributions; this assumption is certainly restrictive, however we think that it is general enough to cover most of the cases of physical interest. First of all we must translate in the probabilistic language the conditions (10) and (11). Let us start by the constraint (11); at this purpose we shall write the following a-priori probability density:

$$(13) \quad P(x) = C_1 e^{-\frac{1}{2E^2} (\mathcal{B} x, \mathcal{B} x)_{R^n}}$$

Similarly we translate in the probabilistic language the condition (10), writing the following conditional probability density of the data vector for a given solution vector^(x):

$$(14) \quad P(h|x) = C_2 e^{-\frac{1}{2\varepsilon^2} ((\mathcal{A} x - h), (\mathcal{A} x - h))_{R^n}}$$

(x) - About the meaning of these probability densities see also ref. (2).

6.

Next we use the following Bayes formula⁽⁷⁾:

$$(15) \quad P(x|h) = \frac{P(x) P(h|x)}{\int P(x) P(h|x) dx}$$

From (13), (14) and (15) we get

$$(16) \quad P(x|h) = C_3 e^{-\frac{1}{2\epsilon^2} \left\{ (x, [a^*a + \lambda^2 \beta^* \beta] x)_{R^n} - 2(a^*h, x)_{R^n} + (h, h)_{R^n} \right\}}$$

where $\lambda = \epsilon/E$, and $(h, h)_{R^n} \equiv \|h\|_Y^2$.

At this point we can find the expected value of the solution random vector \tilde{x} ; we proceed as follows

$$(17) \quad E(\tilde{x}) = C_4 \int x e^{-\frac{1}{2\epsilon^2} \left\{ (x, [a^*a + \lambda^2 \beta^* \beta] x)_{R^n} - 2(a^*h, x)_{R^n} \right\}} dx$$

where the factor $e^{-\frac{1}{2\epsilon^2} (h, h)_{R^n}}$, which does not depend on x , has been included in C_4 . From (17) we see that the mean value of \tilde{x} is given by:

$$(18) \quad \mathcal{C}_x \equiv [a^*a + \lambda^2 \beta^* \beta] x = a^*h$$

in other words it coincides with the solution of the eq. (12): Then we can write the expression of the root-mean-square error; it is given by

$$(19) \quad \sigma = \frac{\epsilon}{\sqrt{n}} \left[\text{Tr}((a^*a + \lambda^2 \beta^* \beta)^{-1}) \right]^{1/2}$$

which is not too far from the stability estimate $M_1(\epsilon, E)$. At this purpose the reader is referred to I where it has been proven that

$$(20) \quad \langle f^1 - f^0 \rangle \leq M_1(\epsilon, E) = \sqrt{2} \epsilon (\mathcal{C}^{-1} y, y)_{R^n}^{1/2}$$

If a^*a and $\beta^* \beta$ have been simultaneously brought to a diagonal form, then the solution (18) and the root-mean-square error (19) will take a very simple form. In fact let us denote with a_j^2 the eigenvalues of a^*a and with b_j^2 the eigenvalues of $\beta^* \beta$ (we can assume that $b_j^2 = 1$ for any j); then the j -th component of the approximating vector x^1 is given by:

$$(21) \quad x_j^1 = \frac{a_j h_j}{(a_j^2 + \lambda^2)}$$

since \mathcal{A}^*h is the vector defined by $(\mathcal{A}^*h)_j = (h, A\psi^j)_Y$; furthermore the root-mean-square error σ is given by:

$$(22) \quad \sigma = \frac{E}{\sqrt{n}} \left(\sum_{j=1}^n \frac{\varepsilon^2}{\varepsilon^2 + (a_j)^2 E^2} \right)^{1/2}.$$

At this point we want to discuss in greater detail the case of simultaneous diagonalization of A^*A and B^*B . Let us return, for the moment, to the general undiscretized problem: If A^*A and B^*B have been simultaneously diagonalized, this implies that it is available for X a basis ψ^1, ψ^2, \dots which is simultaneously orthogonal with respect to both the Y and the Z norms, i. e.

$$(23) \quad (A\psi^i, A\psi^j)_Y = (A^*A\psi^i, \psi^j)_X = a_j^2 \delta_{ij}$$

$$(24) \quad (B\psi^i, B\psi^j)_Z = (B^*B\psi^i, \psi^j)_X = b_j^2 \delta_{ij}$$

In this case the functions $f \in X$ and also h can be expanded in terms of the basis $\{\psi^j\}$ with Fourier coefficients $\{x_j\}$ and $\{h_j\}$ respectively. More precisely we have:

$$(25) \quad f = \sum_{j=1}^{\infty} x_j \psi^j$$

$$(26) \quad h = \underline{h} + \underline{\underline{h}} = \sum_{a_j \neq 0} h_j (A\psi^j) + \underline{\underline{h}}$$

where \underline{h} is in the range of A and $\underline{\underline{h}}$ is in its orthogonal complement. In this situation one can use the methods of partial eigenfunction expansions and the corresponding approximations which have been extensively discussed in I. In order to give an example of the eigenvalues a_j and b_j let us suppose that $\Gamma = \Gamma$ is a circle of radius $0 < a < 1$, while $\partial\Omega$ is, as usual, the unit circle (the analyticity domain Ω is the unit disc); then we have

$$(27) \quad a_j^* \equiv a_j = a^j$$

$$(28) \quad b_j^* \equiv b_j = 1$$

8.

and therefore $a_j/b_j \xrightarrow{j \rightarrow \infty} 0$.

Now we want to try a probabilistic interpretation of the methods of partial eigenfunction expansions. Suppose that the problem has been discretized and let us write the probability densities (13), (14) and (16) as follows

$$(29) \quad P(x) = C_1 e^{-\frac{1}{2} \sum_1^n j \left| \left(\frac{b_j}{E} \right) (x_j - 0) \right|^2}$$

$$(30) \quad P(h|x) = C_2 e^{-\frac{1}{2} \sum_1^n j \left| \left(\frac{a_j}{\varepsilon} \right) (x_j - h_j) \right|^2}$$

$$(31) \quad P(x|h) = C_3 e^{-\frac{1}{2} \left\{ \sum_1^n j \left| \left(\frac{a_j}{\varepsilon} \right) (x_j - h_j) \right|^2 + \sum_1^n j \left| \left(\frac{b_j}{E} \right) (x_j - 0) \right|^2 \right\}}$$

where we have retained, as much as possible, the same notations of the undiscretized problem. Next we assume that the ratios a_j/b_j are decreasing (as the formulae (27) and (28) clearly show in a particular example of the undiscretized case). In this situation one can determine the largest integer α such that $a_\alpha/\varepsilon \geq b_\alpha/E$; it follows that for the first $j \leq \alpha$ components the probability density (30) is more concentrated than the probability density (29); viceversa for the $(\alpha + 1) \leq j \leq n$ components the density (30) is more dispersed than the (29). Observing that the probability density (31) is, essentially, the product of (29) and (30), the considerations above suggest an approximation where the first α components are given by the factors h_j , and the components $(\alpha + 1) \leq j \leq n$ are given by the mean values corresponding to the normal probability density (29), i.e. zero. This approximation can be written as follows

$$(32) \quad h_\alpha = \sum_1^\alpha h_j \psi^j$$

and it coincides with the approximation which is given by the method 1' of the partial eigenfunction expansions of I. It is easy to prove that the approximation (32) is best-possible but for a factor of two (see I), in the sense that

$$(33) \quad \varepsilon_\alpha = \left\| A h_\alpha - h \right\|_Y \leq 2\varepsilon$$

$$(34) \quad E_\alpha = \left\| B h_\alpha \right\|_Z \leq 2E.$$

Finally the reader is referred to I for the various methods which are necessary if one of the numbers ϵ or E is not known.

REFERENCES. -

- (1) - K. Miller and G. A. Viano, J. Math. Phys., to be published.
- (2) - V. F. Turchin, V. P. Kozlov and M. S. Malkevich, Soviet Phys. - Uspekhi 13, 681 (1971).
- (3) - R. R. Cutkosky, Ann. Phys. 54, 350 (1969).
- (4) - J. Kelley, General Topology (Van Nostrand, Princeton, 1955).
- (5) - Z. Nehari, Conformal Mapping (McGraw Hill, New York, 1952).
- (6) - F. John, Comm. on Pure and Applied Mathematics 13, 551 (1960).
- (7) - M. Fisz, Probability Theory and Mathematical Statistics (John Wiley, New York, 1965).

Finally, the reader is referred to I for the various methods which are necessary if one of the numbers r or B is not known.

REFERENCES

- (1) - K. Miller and G. A. Viano, *J. Math. Phys.*, to be published.
- (2) - V. F. Turchin, V. P. Kostov and M. S. Mal'nev, *Soviet Phys. Technol. Phys.*, 13, 681 (1977).
- (3) - H. R. Cuntz, *Ann. Phys.*, 54, 350 (1969).
- (4) - J. Keller, *Geometrical Topology* (Van Nostrand, Reinhold, 1970).
- (5) - X. Robert, *Geometrical Mapping* (McGraw-Hill, New York, 1969).
- (6) - F. John, *Comm. on Pure and Applied Mathematics*, 12, 151 (1958).
- (7) - M. Tjur, *Proximity Theory and Mathematical Statistics* (Wiley, New York, 1965).