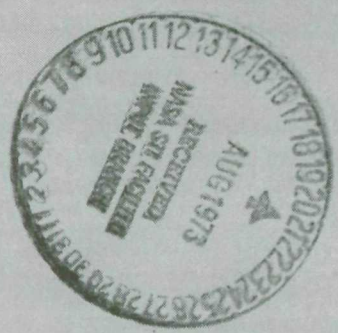
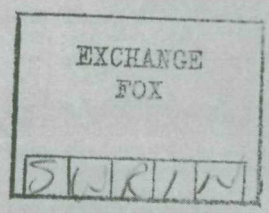
FOA REPORTS

NONITERATIVE ESTIMATION OF A NONLINEAR PARAMETER

Arne Bergström





FÖRSVARETS FORSKNINGSANSTALT · (FOA)
RESEARCH INSTITUTE OF NATIONAL DEFENCE
STOCKHOLM 80 · SWEDEN

FOA Reports Vol. 7 No. 3 pp. 1–8 April 1973

FOA REPORTS

FOA REPORTS is a monograph series of unclassified reports dealing with research work of general interest carried out by the staff and consultants of the Research Institute of National Defence (FOA). Documents appearing in FOA REPORTS express the personal opinions of the authors and do not necessarily represent the official view on the subject.

FOA REPORTS appears at irregular intervals, and each issue is priced separately. Orders for single issues or a year's subscription should be sent to the Editorial Office (see below).

Editorial Office: FOA P, 104 50 Stockholm 80, Sweden
Tel. 08-631500 Cables: Foageneral, Stockholm

FOA Reports, Vol. 7, No. 3, 1973

NONITERATIVE ESTIMATION OF A NONLINEAR PARAMETER*

ARNE BERGSTRÖM

ABSTRACT

An algorithm is described which solves the parameters $\mathbf{x}=(x_1, x_2, ..., x_m)$ and p in an approximation problem $\mathbf{Ax}\approx\mathbf{y}(p)$, where the parameter p occurs nonlinearly in $\mathbf{y}$. Instead of linearization methods, which require an approximate value of p to be supplied as *a priori* information, and which may lead to the finding of local minima, the proposed algorithm finds the global minimum by permitting the use of series expansions of arbitrary order, exploiting an *a priori* knowledge that the addition of a particular function, corresponding to a new column in $\mathbf{A}$, will not improve the goodness of the approximation.

I. INTRODUCTION

In the present paper the following approximation problem will be studied. Given

(1) a function $y(s, p)$ of an independent variable s (the extension to functions of several variables is trivial) and containing a parameter p which occurs nonlinearly, and

(2) a set of functions $f_1(s)$, $f_2(s)$, ..., $f_m(s)$ of s,

determine a linear combination $g(s, x)=x_1 f_1(s)+x_2 f_2(s)+...+x_m f_m(s)$ of the functions in (2), and a value of the parameter p in (1), such that the residual function $\varepsilon(s)=g(s, x)-y(s, p)$ over a certain interval in s is minimized in some sense.

In discrete formalism, where the functions are given at discrete points $s_1, s_2, ..., s_n$, the approximation problem may for a large class of minimizing criteria be expressed as

$$\mathbf{Ax}=\mathbf{y}(p)+\boldsymbol{\epsilon}, \tag{1a}$$

$$\mathbf{B}\boldsymbol{\epsilon}=0, \tag{1b}$$

where $\mathbf{A}$ is a $n\times m$ matrix with elements $a_{ij}=f_j(s_i)$,

$$\mathbf{x} \text{ is the column vector } \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix},$$

$$\mathbf{y}(p) \text{ is the column vector } \begin{pmatrix} y(s_1, p) \\ y(s_2, p) \\ \vdots \\ y(s_n, p) \end{pmatrix},$$

and $\mathbf{B}$ is a matrix which is derived from the criterion used for minimizing the residuals

$$\boldsymbol{\epsilon}=\begin{pmatrix} \varepsilon(s_1) \\ \varepsilon(s_2) \\ \vdots \\ \varepsilon(s_n) \end{pmatrix}.$$

The most extensively used minimizing criterion is the least squares criterion, in which $\boldsymbol{\epsilon}^T\boldsymbol{\epsilon}$ is minimized. As is easily seen by differentiation, this criterion leads to a matrix $\mathbf{B}_q$ equal to $\mathbf{A}^T$, the transposed coefficient matrix, with an additional row due to the nonlinear parameter and containing elements

$$(b_q)_{m+1,i}=-\frac{\partial y(s_i, p)}{\partial p}.$$

Another example of a minimizing criterion is moment matching, where the first $m+1$ moments of the right and left sides of Eqn. (1a) are set equal. In the above formalism this is accomplished by using a matrix

$$\mathbf{B}_m=\begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ s_1 & s_2 & s_3 & \dots & s_n \\ s_1^2 & s_2^2 & s_3^2 & \dots & s_n^2 \\ s_1^3 & s_2^3 & s_3^3 & \dots & s_n^3 \\ \vdots & \vdots & \vdots & & \vdots \\ s_1^m & s_2^m & s_3^m & \dots & s_n^m \end{pmatrix}.$$

The approximation problem may be illustrated by the following geometrical picture, see Fig. 1. We represent a

* This work was in part performed at Southern Methodist University, Dallas, Texas, on NASA Grant NsG 708.

ISBN-91-7056-023-4

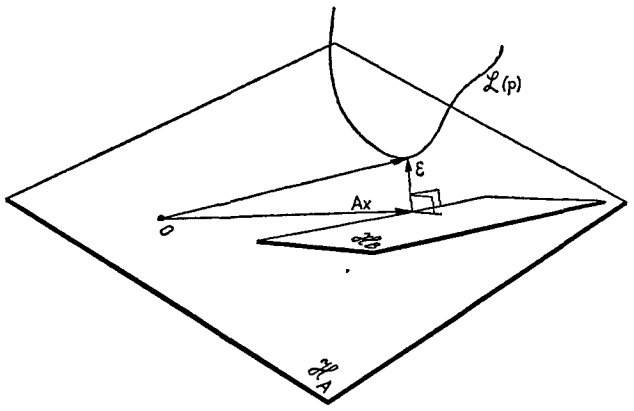


Fig. 1. Geometrical representation of the nonlinear approximation problem.

function $\varphi(s)$, given at the discrete points $s_1, s_2, \ldots, s_n$ by a vector $\varphi = \begin{pmatrix} \varphi(s_1) \\ \varphi(s_2) \\ \vdots \\ \varphi(s_n) \end{pmatrix}$, as a point in a n-dimensional space. As p varies, the vector $\mathbf{y}(p)$ discussed above will in this representation describe a curve $\mathcal{L}(p)$; similarly, the left-hand side of Eqn. (1a) will be represented by a m-dimensional hyperplane $\mathcal{H}_A$ spanned by the columns of the matrix $\mathbf{A}$, and the vector $\boldsymbol{\epsilon}$ will connect a point on $\mathcal{L}(p)$ with a point on $\mathcal{H}_A$. The criterion quantity which is used to measure the goodness of the approximation, or in general how "close" two functions are to each other, defines in this picture a metric of the n-dimensional space. In the class of minimizing criteria expressed above as Eqn. (1) the metric is $\boldsymbol{\epsilon}^T\mathbf{M}\boldsymbol{\epsilon}$ with $\mathbf{A}^T\mathbf{M}=\mathbf{B}$. This metric corresponds to the ordinary Euclidean metric $\boldsymbol{\epsilon}^T\boldsymbol{\epsilon}$ after an affine transformation characterized by a $n \times n$ transformation matrix $\mathbf{N}$ such that $\mathbf{N}^T\mathbf{N}=\mathbf{M}$, and is expressed in differential form by Eqn. (1b) with the implication that the shortest length of the residual vector $\boldsymbol{\epsilon}$ in the sense of the criterion used is when it is perpendicular to the hyperplane $\mathcal{H}_B$ which is spanned by the rows of $\mathbf{B}$.

II. PARAMETER SEARCH AND TAYLOR LINEARIZATION

For the linear approximation problem of determining $\mathbf{x}$ for a fixed value of p from Eqn. (1), there exists an explicit and, if $(\mathbf{BA})^{-1}$ is nonsingular, unique solution, $\mathbf{x}=(\mathbf{BA})^{-1}\mathbf{By}$. [To insure numerical stability, a somewhat different computational approach is, however, in general required (Householder, 1953; see also Golub, 1965).] The nonlinear part of the problem, i.e. the determination of the nonlinearly occurring parameter p, constitutes a somewhat more intricate problem both because there exists no such simple algorithm as in the linear case and because there may be several minima for the criterion quantity and ways must be found to ascertain that actually the lowest one is found. For this problem two classes of methods, employing parameter search or Taylor linearization, are usually used, either separately or in combination, and as a background for the following chapter a brief outline of these methods will here be given.

In the parameter search, the linear approximation problem is solved for different fixed values of the nonlinear parameter, and the corresponding measures of the goodness of the approximation are computed. In this way one aims at obtaining a coarse picture of the criterion quantity as a function of the nonlinear parameter in order to discriminate between the region where the absolute minimum lies and regions where possible local minima may be found. In the former region a successively finer subdivision can then be made until the minimum is known within a required accuracy.

In the linearization methods, a Taylor expansion of $\mathbf{y}(p)$ to the first order in p is performed around a value p_0 (which is assumed to be sufficiently close to the exact value), $\mathbf{y}(p)=\mathbf{y}(p_0)+\mathbf{y}'(p_0)\,\Delta p+\mathrm{O}(\Delta p^2)$, where

$$\mathbf{y}'(p_0)=\begin{bmatrix} \dfrac{\partial y(s_1,p_0)}{\partial p} \\ \dfrac{\partial y(s_2,p_0)}{\partial p} \\ \vdots \\ \dfrac{\partial y(s_n,p_0)}{\partial p} \end{bmatrix}.$$

If we let $\mathbf{A}'$ denote a matrix which consists of the coefficient matrix $\mathbf{A}$ as above with the additional column $-\mathbf{y}'(p_0)$, a first approximation of the vector $\mathbf{x}'=\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \\ \Delta p \end{pmatrix}$ is obtained as $\mathbf{x}'=(\mathbf{BA}')^{-1}\mathbf{By}(p_0)$, where $\mathbf{B}=(\mathbf{A}')^T\mathbf{M}$. A new linearization and solution is then made around $p_0^{(1)}=p_0+\Delta p^{(0)}$, and the procedure is repeated until Δp lies within the required accuracy.

Of the two classes of methods, the Taylor linearization is the fastest. A restriction is, however, that the initial guess must be sufficiently close to the exact value, otherwise there is always the risk that the method will find a local minimum. At the cost of a substantially more tedious computational procedure, the parameter search to a great extent excludes the possibility of finding local minima. However, if a too coarse scanning is made, there is also in this case a small risk that the absolute minimum may be lost.

To illustrate the above difficulties we will study the following approximation problem.

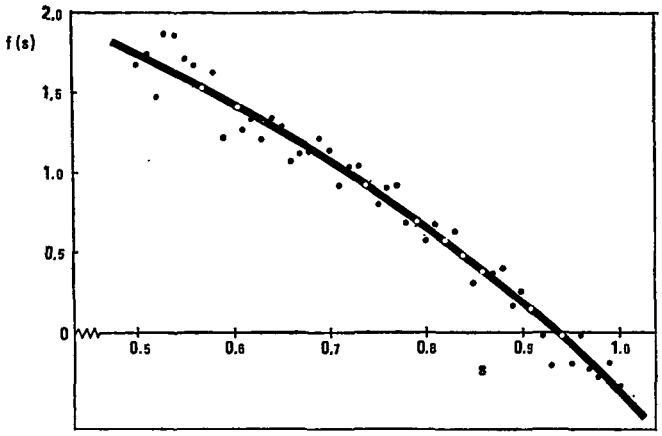


Fig. 2. The function $f(s) = 3 + s - \exp(3s/2) + \Phi(0.9 \sin(3s/2), 0.1)$ on the interval $0.5 \leq s \leq 1.0$. Solid line represents least-squares fit.

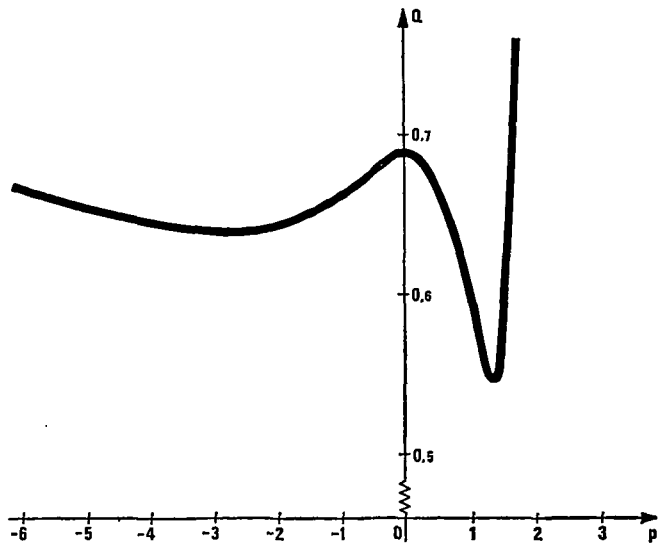


Fig. 3. Sum of squares of residuals Q *versus* the value of the nonlinear parameter p. (Note the suppressed zero.)

Example: Determine the parameters x_1, x_2 and p in the expression

$$g(s) = x_1 + x_2 s - e^{ps}$$

to give the best approximation in the least-squares sense to the points given in Fig. 2. (These points are computed as

$$f(s) = 4 + s - e^{\frac{1}{2}s} + \Phi(-0.9 \sin \tfrac{3}{2}s, 0.1),$$

where $\Phi(m, \sigma)$ denotes a gaussian distribution with mean m and standard deviation σ, which is included in order to simulate a disturbance with a systematic as well as a stochastic component.)

Results: As is seen in Fig. 2, the function $f(s)$ has a main trend which is convex from above. As long as p is sufficiently different from zero, this is also the case for the approximating function $g(s)$, both if p is negative and if p is positive. When p becomes closer to zero, however, the approximating function degenerates to a linear function, and it is to be expected that the possibility to obtain a good fit to $f(s)$ is less than with a convex function. This explains the appearance of a local minimum at $p \approx -2.5$ in addition to the lowest minimum at $p \approx 1.5$ in Fig. 3, where the criterion quantity Q (the sum of squares of the residuals) for the best choice of the corresponding linear parameter is plotted against the values of the nonlinear parameter p. This structure of the functional dependence of the criterion quantity may lead to difficulties when attempting to use the two algorithms discussed above to compute the nonlinear parameter. If a negative value of p is used as starting value in the linearization method, the iterative procedure will in general find the local minimum instead. This turns out to be true also for more elaborate linearization methods such as those proposed by Davidon (1959) and Powell (1965). Due to the sharpness of the lowest minimum there may be some danger that also the parameter search misses this minimum if a too coarse initial scanning is performed.

In the new, noniterative algorithm which will be discussed in the following, these difficulties are removed.

III. THE NEW ALGORITHM

In contrast to the Taylor linearization method, which requires an approximate value of the nonlinear parameter as *a priori* information, the new algorithm which we will develop here exploits a possible *a priori* knowledge that the addition of a particular function to the left-hand side of Eqn. (1a) will not improve the goodness of the approximation, i.e. that the vector in the above geometrical model which corresponds to the additional function is orthogonal, in the sense of the criterion, to the residual vector ϵ. Especially if the dimensionality of the space is high, as is desired for reasons of numerical accuracy, there is a multitude of functions obeying this orthogonality requirement and, as we shall see in Chapter IV, experience also shows that there is as a rule no difficulty in finding such functions; suitable functions are even in many cases suggested by the very assumptions made in formulating the problem, e.g. that certain terms can be neglected in the mathematical description of the actual problem.

Using the picture of Chapter I, the original approximation problem corresponds to the geometrical problem of finding a point on the curve $\mathcal{L}(p)$ and a point on the hyperplane $\mathcal{H}_A$ such that the residual vector ϵ, orthogonal to $\mathcal{H}_B$, is minimized. By including a new function (as column $m+1$) in A as discussed above, a new hyperplane $\mathcal{H}'_A$, now of dimension $m+1$, is formed (Fig. 4). Due to the assumptions inherent in the choice of the additional function, this hyperplane has the property of being

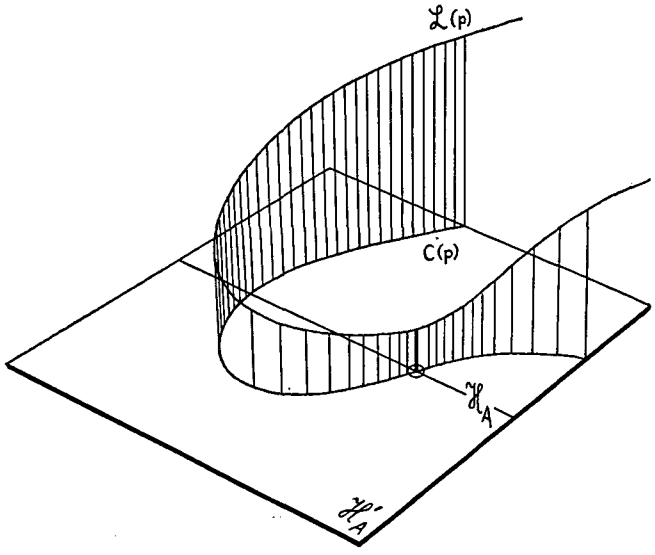


Fig. 4. The nonlinear parameter is solved from the best intersection between the hyperplane $\mathcal{H}_A$ (in the figure one-dimensional) and the projection of $c(p)$ on the hyperplane $\mathcal{H}'_A$.

orthogonal in the sense of the metric to the residual vector ϵ above.

The solution to Eqn. (1), $\mathbf{x}=(\mathbf{BA})^{-1}\mathbf{B}\mathbf{y}(p)$, for brevity called $\mathbf{c}(p)$ in the following, gives $\mathbf{x}$ as a function of p corresponding to a curve on the hyperplane $\mathcal{H}'_A$ which is the projection of $\mathcal{L}(p)$. (We here assume that the vector corresponding to the additional function is not orthogonal to $\mathcal{L}(p)$.) As follows immediately from the above constructions, one of the intersections between the projected curve and the original hyperplane $\mathcal{H}_A$ is identical with the endpoint of the residual vector $\boldsymbol{\epsilon}$, which we wanted to determine. This implies that the value of the nonlinear parameter can be solved by setting the coefficient x_{m+1} for the additional function equal to zero, i.e. the nonlinear parameter is one of the roots to the equation $c_{m+1}(p)=0$. This expression is greatly simplified if $\mathbf{y}(p)$ is given in the form of a series expansion.

If the functions $\mathbf{y}(p)$ are given as power series expansions, assuming the power series expansion of $c_{m+1}(p)$ to be convergent, the new algorithm can be expressed in algorithms for the solution of linear approximation problems and polynomial roots by using the following lemma.

LEMMA. Let a linear approximation problem with n observations and m unknown parameters be given as

$$\mathbf{Ax}=\mathbf{y}+\boldsymbol{\epsilon},$$
$$\mathbf{B}\boldsymbol{\epsilon}=0,$$

where $(\mathbf{BA})^{-1}$ is nonsingular, so that there exists a unique solution. If we perform an arbitrary decomposition of $\mathbf{y}$ into $\mathbf{y}=\mathbf{y}_1+\mathbf{y}_2+\ldots+\mathbf{y}_l$, the solution $\mathbf{x}$ to the original problem may be written as $\mathbf{x}=\mathbf{x}_1+\mathbf{x}_2+\ldots+\mathbf{x}_l$, where

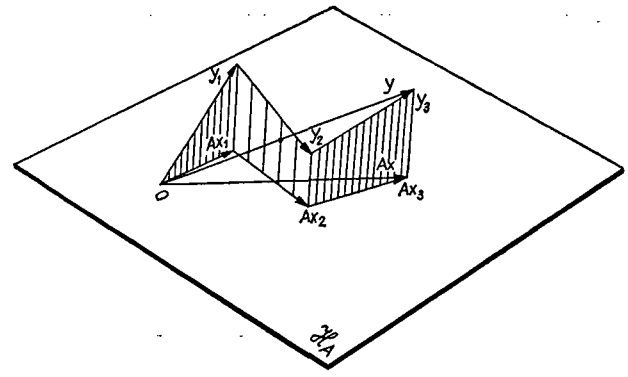


Fig. 5. Geometrical interpretation of the decomposition lemma.

$\mathbf{x}_i$, $i=1, 2, \ldots, l$, are the solutions $\mathbf{x}_i=(\mathbf{BA})^{-1}\mathbf{B}\mathbf{y}_i$ to the decomposed approximation problems

$$\mathbf{Ax}_i=\mathbf{y}_i+\boldsymbol{\epsilon}_i,$$
$$\mathbf{B}\boldsymbol{\epsilon}_i=0.$$

Proof. In each of the decomposed approximation problems, $\mathbf{Ax}_i=\mathbf{y}_i+\boldsymbol{\epsilon}_i$, $\mathbf{B}\boldsymbol{\epsilon}_i=0$, we have $n+m$ equations which, since $(\mathbf{BA})^{-1}$ is assumed to be nonsingular, completely determine the $n+m$ unknowns $\mathbf{x}_i$ and $\boldsymbol{\epsilon}_i$. By summing over i, $i=1, 2, \ldots l$, we obtain

$$\mathbf{A}(\mathbf{x}_1+\mathbf{x}_2+\ldots+\mathbf{x}_l)=\mathbf{y}_1+\mathbf{y}_2+\ldots+\mathbf{y}_l+\boldsymbol{\epsilon}_1+\boldsymbol{\epsilon}_2+\ldots+\boldsymbol{\epsilon}_l,$$
$$\mathbf{B}(\boldsymbol{\epsilon}_1+\boldsymbol{\epsilon}_2+\ldots+\boldsymbol{\epsilon}_l)=0,$$

and, since $\mathbf{y}_1+\mathbf{y}_2+\ldots+\mathbf{y}_l=\mathbf{y}$, we obtain

$$\mathbf{x}=\mathbf{x}_1+\mathbf{x}_2+\ldots+\mathbf{x}_l. \quad \text{Q.E.D.}$$

In the geometrical picture described in Chapter I this lemma has a simple interpretation, see Fig. 5. The vector $\mathbf{Ax}$ is the projection of the vector $\mathbf{y}$ on the hyperplane $\mathcal{H}_A$. The lemma merely states that when $\mathbf{y}$ is expressed as a sum of vectors $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_l$, the vector $\mathbf{Ax}$ may be expressed as the sum of the projections of $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_l$ on $\mathcal{H}_A$, i.e. the vectors $\mathbf{Ax}_1, \mathbf{Ax}_2, \ldots, \mathbf{Ax}_l$.

Returning now to the nonlinear approximation problem discussed earlier,

$$\mathbf{Ax}=\mathbf{y}(p)+\boldsymbol{\epsilon},$$
$$\mathbf{B}\boldsymbol{\epsilon}=0,$$

where the coefficient for the artificial dimension introduced is x_{m+1}, we write the first system of equations in explicit form

$$\begin{cases} a_{11}x_1+a_{12}x_2+\ldots+a_{1,m+1}x_{m+1}=y_{10}+y_{11}p+y_{12}p^2+\ldots+\varepsilon_1 \\ a_{21}x_1+a_{22}x_2+\ldots+a_{2,m+1}x_{m+1}=y_{20}+y_{21}p+y_{22}p^2+\ldots+\varepsilon_2 \\ \vdots \\ a_{n1}x_1+a_{n2}x_2+\ldots+a_{n,m+1}x_{m+1}=y_{n0}+y_{n1}p+y_{n2}p^2+\ldots \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad +\varepsilon_n. \end{cases}$$

In correspondence with the power series expansion in the

right-hand side, we now decompose $x_1, x_2, x_3, \ldots, x_{m+1}$ as $x_1 = x_{10} + x_{11} + x_{12} + \ldots$, $x_2 = x_{20} + x_{21} + x_{22} + \ldots$, etc. After introduction of the new unknowns $x'_{10} = x_{10}$, $x'_{11} = x_{11}/p$, $x'_{12} = x_{12}/p^2, \ldots, x'_{20} = x_{20}$, $x'_{21} = x_{21}/p$, $x'_{22} = x_{22}/p^2, \ldots$, etc., we may from the overdetermined equation system above form a new set of overdetermined equation systems:

$$\begin{cases} a_{11}x'_{10} + a_{12}x'_{20} + \ldots + a_{1,m+1}x'_{m+1,0} = y_{10} + \varepsilon_{10} \\ a_{21}x'_{10} + a_{22}x'_{20} + \ldots + a_{2,m+1}x'_{m+1,0} = y_{20} + \varepsilon_{20} \\ \vdots \\ a_{n1}x'_{10} + a_{n2}x'_{20} + \ldots + a_{n,m+1}x'_{m+1,0} = y_{n0} + \varepsilon_{n0}, \end{cases}$$

from which the unknowns $\mathbf{x}'_0 = (x'_{10}, x'_{20}, \ldots, x'_{m+1,0})$ can be determined;

$$\begin{cases} a_{11}x'_{11}p + a_{12}x'_{21}p + \ldots + a_{1,m+1}x'_{m+1,1}p = y_{11}p + \varepsilon_{11} \\ a_{21}x'_{11}p + a_{22}x'_{21}p + \ldots + a_{2,m+1}x'_{m+1,1}p = y_{21}p + \varepsilon_{21} \\ \vdots \\ a_{n1}x'_{11}p + a_{n2}x'_{21}p + \ldots + a_{n,m+1}x'_{m+1,1}p = y_{n1}p + \varepsilon_{n1}, \end{cases}$$

from which, after division by p, the unknowns $\mathbf{x}'_1 = (x'_{11}, x'_{21}, \ldots, x'_{m+1,1})$ can be determined (since all residuals $\varepsilon_{11}, \varepsilon_{21}, \ldots, \varepsilon_{n1}$ are affected equally by the divison with p, the division does not change the problem);

$$\begin{cases} a_{11}x'_{12}p^2 + a_{12}x'_{22}p^2 + \ldots + a_{1,m+1}x'_{m+1,2}p^2 = y_{12}p^2 + \varepsilon_{12} \\ a_{21}x'_{12}p^2 + a_{22}x'_{22}p^2 + \ldots + a_{2,m+1}x'_{m+1,2}p^2 = y_{22}p^2 + \varepsilon_{22} \\ \vdots \\ a_{n1}x'_{12}p^2 + a_{n2}x'_{22}p^2 + \ldots + a_{n,m+1}x'_{m+1,2}p^2 = y_{n2}p^2 + \varepsilon_{n2}, \end{cases}$$

from which after division by p^2 the unknowns $\mathbf{x}'_2 = (x'_{12}, x'_{22}, \ldots, x'_{m+1,2})$ can be determined;

etc. for the other powers of p.

Since only the right-hand side differs between the cases, the solution of these equation systems is a very fast procedure.

According to the decomposition lemma the solution vector $\mathbf{x}$ to the original problem is given as

$$\mathbf{x} = \mathbf{x}_0 + \mathbf{x}_1 + \mathbf{x}_2 + \ldots = \mathbf{x}'_0 + \mathbf{x}'_1 p + \mathbf{x}'_2 p^2 + \ldots .$$

Following the discussion in the beginning of this chapter, the value of the nonlinear parameter can then be solved from the equation $x_{m+1} = 0$, selecting the best root in the sense of the criterion used. Once the value of p is known, $x_1, x_2, \ldots, x_m$ can also be computed as $x_1 = x'_{10} + x'_{11}p + x'_{12}p^2 + \ldots$, $x_2 = x'_{20} + x'_{21}p + x'_{22}p^2 + \ldots$, etc.

To illustrate the conciseness with which the new algorithm can be programmed using existing algorithms for linear least-squares problems and polynomial roots, and to provide a model for its implementation in other programming languages, a FORTRAN program of a least-squares version expressed in subroutines from the IBM Scientific Subroutine Package (IBM, 1970) is given in Fig. 6.

An alternative to the above geometrical approach of describing the algorithm is as follows.

The *a priori* assumption is equivalent to saying that the approximation problem

$$(\mathbf{A} \quad \mathbf{a}_{m+1}) \begin{pmatrix} \mathbf{x} \\ x_{m+1} \end{pmatrix} = \mathbf{y}(p) + \boldsymbol{\varepsilon},$$

$$\begin{pmatrix} \mathbf{B} \\ \mathbf{b}^T_{m+1} \end{pmatrix} \boldsymbol{\varepsilon} = 0,$$

has the solution $x_{m+1} = 0$. Eliminating $\boldsymbol{\varepsilon}$ we obtain

$$\begin{pmatrix} \mathbf{BA} & \mathbf{Ba}_{m+1} \\ \mathbf{b}^T_{m+1}\mathbf{A} & \mathbf{b}^T_{m+1}\mathbf{a}_{m+1} \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ x_{m+1} \end{pmatrix} = \begin{pmatrix} \mathbf{B} \\ \mathbf{b}^T_{m+1} \end{pmatrix} \mathbf{y}(p).$$

Assuming the matrix $(\mathbf{BA})$ to be nonsingular, x_{m+1} can be solved from this equation, e.g. by the method of bordering (Faddeeva, 1959), i.e.

$$x_{m+1} = \frac{1}{\alpha_m} \mathbf{b}^T_{m+1}(\mathbf{I} - \mathbf{P}_{A/B})\,\mathbf{y}(p), \qquad \alpha_m = \mathbf{b}^T_{m+1}(\mathbf{I} - \mathbf{P}_{A/B})\mathbf{a}_{m+1},$$

where

$$\mathbf{P}_{A/B} = \mathbf{A}(\mathbf{BA})^{-1}\mathbf{B}$$

is a projection operator, and we have assumed that $\alpha_m \neq 0$. If we define $\mathbf{r}_{m+1}$ by

$$\mathbf{r}_{m+1} = (\mathbf{I} - \mathbf{P}^T_{A/B})\,\mathbf{b}_{m+1}$$

(note that $\mathbf{r}_{m+1}$ is easily computed as the residual vector corresponding to the approximation problem with right-hand side $\mathbf{b}_{m+1}$), we get the simple formula

$$x_{m+1} = \frac{1}{\alpha_m} \mathbf{r}^T_{m+1}\,\mathbf{y}(p), \qquad \alpha_m = \mathbf{r}^T_{m+1}\,\mathbf{a}_{m+1}.$$

The condition $x_{m+1} = 0$ now gives a nonlinear equation from which p can be solved. In particular, if $\mathbf{y}(p)$ is given by a power series expansion $\mathbf{y}(p) = \mathbf{y}_0 + \mathbf{y}_1 p + \mathbf{y}_2 p^2 + \ldots$, we get $a_0 + a_1 p + a_2 p^2 + \ldots = 0$, where $a_k = \mathbf{r}^T_{m+1}\mathbf{y}_k$.

In the case of the least-squares criterion, $\mathbf{P}_{A/B}$ is symmetric, and the formula for $\mathbf{r}_{m+1}$ reduces to

$$\mathbf{r}_{m+1} = [\mathbf{I} - \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T]\,\mathbf{b}_{m+1}.$$

As above, $\mathbf{r}_{m+1}$ is here easily computed (as the residual vector corresponding to the least-squares problem with right-hand side $\mathbf{b}_{m+1}$).

The method can be characterized by saying that an *a priori* vector $\mathbf{a}_{m+1}$ is used instead of $-\partial\mathbf{y}/\partial p$ in the $(m+1)$-th column in $\mathbf{A}$. To guard against an unfortunate choice of the artificial dimension, one can after the first value $p = p^*$ has been obtained choose $\hat{\mathbf{a}}_{m+1} = -(\partial\mathbf{y}/\partial p)_{p=p^*}$ as new vector and recompute p. This method can of course easily be iterated, and since only the right-hand side changes, it is a very fast procedure.

We will here also give the expressions for the statistical uncertainties in the estimates of the parameters under

```
C                                                                       TGTH  10
C     ..................................................................TGTH  20
C                                                                       TGTH  30
C        SUBROUTINE TGTHPL                                              TGTH  40
C                                                                       TGTH  50
C        PURPOSE                                                        TGTH  60
C           LEAST SQUARES SOLUTION OF THE VECTOR X AND THE PARAMETER P  TGTH  70
C           IN AN OVERDETERMINED EQUATION SYSTEM WRITTEN IN MATRIX FORM TGTH  80
C           AS A*X=Y(P), WHERE P OCCURS NON-LINEARLY IN THE VECTOR Y(P) TGTH  90
C                                                                       TGTH 100
C        USAGE                                                          TGTH 110
C           CALL TGTHPL(A,V,S,X,R,AX,Y,Q,E,W1,W2,N,M,L,IER1,IER2,IER3)  TGTH 120
C                                                                       TGTH 130
C        DESCRIPTION OF PARAMETERS                                      TGTH 140
C           A      - THE N BY M INPUT MATRIX OF COEFFICIENTS FOR THE    TGTH 150
C                    LINEAR UNKNOWNS. THE M-TH COLUMN REPRESENTS THE    TGTH 160
C                    ARTIFICIAL DIMENSION USED TO FORM THE TANGENT      TGTH 170
C                    HYPERPLANE, SEE UNDER METHOD BELOW                 TGTH 180
C           V      - THE N BY L INPUT MATRIX OF COEFFICIENTS IN THE     TGTH 190
C                    TAYLOR EXPANSION OF THE RIGHT HAND SIDE IN POWERS  TGTH 200
C                    OF THE NON-LINEAR UNKNOWN, ORDERED FROM LOW TO     TGTH 210
C                    HIGH ORDER                                         TGTH 220
C           S      - OUTPUT VARIABLE CONTAINING THE SUM OF SQUARES OF   TGTH 230
C                    THE RESIDUALS                                      TGTH 240
C           X      - OUTPUT VECTOR OF LENGTH M CONTAINING THE SOLUTION. TGTH 250
C                    LOCATION M CONTAINS THE NON-LINEAR UNKNOWN         TGTH 260
C           R      - CONTAINS ON RETURN THE M BY L MATRIX OF COEFFI-     TGTH 270
C                    CIENTS IN THE EXPANSION OF THE LINEAR UNKNOWNS     TGTH 280
C           AX     - WORKING STORAGE OF LENGTH MAX(2*N,L). THE FIRST    TGTH 290
C                    N LOCATIONS CONTAINS ON RETURN THE LEFT HAND SIDE  TGTH 300
C                    VECTOR A*X                                         TGTH 310
C           Y      - CONTAINS ON RETURN THE RIGHT HAND SIDE VECTOR Y    TGTH 320
C                    OF LENGTH N                                        TGTH 330
C           Q      - VECTOR OF LENGTH L-1 CONTAINING ON RETURN THE REAL TGTH 340
C                    PARTS OF THE ROOTS TO THE POLYNOMIAL EQUATION FOR  TGTH 350
C                    THE NON-LINEAR UNKNOWN                             TGTH 360
C           E      - VECTOR OF LENGTH L-1 CONTAINING ON RETURN THE      TGTH 370
C                    IMAGINARY PARTS OF THE ROOTS TO THE POLYNOMIAL     TGTH 380
C                    EQUATION FOR THE NON-LINEAR UNKNOWN                TGTH 390
C           W1     - WORKING STORAGE OF LENGTH MAX(N*M,L)               TGTH 400
C           W2     - WORKING STORAGE OF LENGTH N*L                      TGTH 410
C           N      - NUMBER OF ROWS IN EQUATION SYSTEM                  TGTH 420
C           M      - NUMBER OF COLUMNS IN EQUATION SYSTEM (INCLUDING    TGTH 430
C                    ARTIFICIAL DIMENSION)                              TGTH 440
C           L      - NUMBER OF TERMS IN TAYLOR EXPANSION OF Y(P)        TGTH 450
C           IER1   - ERROR MESSAGE FROM LINEAR LEAST SQUARES SUBROUTINE TGTH 460
C           IER2   - ERROR MESSAGE FROM POLYNOMIAL ROOTS SUBROUTINE     TGTH 470
C           IER3   - ERROR MESSAGE FROM TGTHPL                          TGTH 480
C                    IER3=0 - NO ERROR                                  TGTH 490
C                    IER3=1 - NO REAL ROOTS. THE RESULTS GIVEN IN       TGTH 500
C                             THIS CASE ARE COMPUTED USING THE BEST     TGTH 510
C                             REAL PART OF THE ROOTS                    TGTH 520
C                                                                       TGTH 530
C        REMARKS                                                        TGTH 540
C           MATRICES A,V,R ARE GENERAL MATRICES                         TGTH 550
C                                                                       TGTH 560
C        SUBROUTINES AND FUNCTION SUBPROGRAMS REQUIRED                  TGTH 570
C           IBM SYSTEM/360 SCIENTIFIC SUBROUTINE PACKAGE ROUTINES:      TGTH 580
C           MCPY,RCPY,GMSUB,GMPRD,MATA,LLSQ,PVAL,POLRT                  TGTH 590
C           (REFERENCE: IBM PUBLICATION GH20-0205-4)                    TGTH 600
C                                                                       TGTH 610
C        METHOD                                                         TGTH 620
C           THE ROUTINE USES THE NON-ITERATIVE ALGORITHM DESCRIBED BY   TGTH 630
C           A. BERGSTROM, FOA REPORT B4059-M4 (1973).                   TGTH 640
C           A VECTOR CORRESPONDING TO A FUNCTION WHICH WILL NOT IMPROVE TGTH 650
C           THE GOODNESS OF THE APPROXIMATION HAS TO BE SUPPLIED BY     TGTH 660
C           THE USER AS COLUMN M IN MATRIX A. THIS VECTOR IS USED       TGTH 670
C           TO FORM A TANGENT HYPERPLANE UPON WHICH THE CURVE Y(P) IS   TGTH 680
C           PROJECTED. THE SOLUTION TO THE ORIGINAL PROBLEM IS THEN     TGTH 690
C           OBTAINED BY SETTING THE COEFFICIENT FOR THE ADDITIONAL      TGTH 700
C           VECTOR EQUAL TO ZERO                                        TGTH 710
C                                                                       TGTH 720
C     ..................................................................TGTH 730
C                                                                       TGTH 740
```

Fig. 6. FORTRAN IV program print-out of least-squares version of the algorithm.

```
      SUBROUTINE TGTHPL(A,V,S,X,R,AX,Y,Q,E,W1,W2,N,M,L,IER1,IER2,IER3)  TGTH 750
      DIMENSION A(1),V(1),X(1),R(1),AX(1),Y(1),Q(1),E(1),W1(1),W2(1)   TGTH 760
      CALL MCPY(A,W1,N,M,0)                                            TGTH 770
      CALL MCPY(V,W2,N,L,0)                                            TGTH 780
      CALL LLSQ(W1,W2,N,M,L,R,Y,0.,IER1,AX)                            TGTH 790
      CALL RCPY(R,M,W1,M,L,0)                                          TGTH 800
      IR=L-1                                                           TGTH 810
      CALL POLRT(W1,W2,IR,Q,E,IER2)                                    TGTH 811
      IER3=1                                                           TGTH 820
      DO 10 I=1,IR                                                     TGTH 830
      CALL EVAL(I,Q,R,X,V,Y,A,AX,S,N,M,L,W1,W2)                        TGTH 840
      IF(I.NE.1.AND.(IER3.NE.1.AND.(E(I).NE.0..OR.S.GE.S0).OR.         TGTH 850
     &   E(I).NE.0..AND.S.GE.S0)) GO TO 10                             TGTH 860
      IF(E(I).EQ.0.)IER3=0                                             TGTH 870
      I0=I                                                             TGTH 880
      S0=S                                                             TGTH 890
   10 CONTINUE                                                         TGTH 900
      CALL EVAL(I0,Q,R,X,V,Y,A,AX,S,N,M,L,W1,W2)                       TGTH 910
      X(M)=Q(I0)                                                       TGTH 920
      RETURN                                                           TGTH 930
      END                                                              TGTH 940
      SUBROUTINE EVAL(I,Q,R,X,V,Y,A,AX,S,N,M,L,W1,W2)                  TGTH 950
      DIMENSION Q(1),R(1),X(1),V(1),Y(1),A(1),AX(1),W1(1),W2(1)        TGTH 960
      DO 10 J=1,M                                                      TGTH 970
      CALL RCPY(R,J,W1,M,L,0)                                          TGTH 980
   10 CALL PVAL(X(J),Q(I),W1,L)                                        TGTH 990
      DO 20 J=1,N                                                      TGTH1000
      CALL RCPY(V,J,W2,N,L,0)                                          TGTH1010
   20 CALL PVAL(Y(J),Q(I),W2,L)                                        TGTH1020
      CALL GMPRD(A,X,AX,N,M,1)                                         TGTH1030
      CALL GMSUB(Y,AX,W1,N,1)                                          TGTH1040
      CALL MATA(W1,S,N,1,0)                                            TGTH1050
      RETURN                                                           TGTH1060
      END                                                              TGTH1070
```

Fig. 6 (continued).

the assumption that the ε_i are independent, have zero mean, and standard deviation σ. As discussed above, the solution to the approximation problem is given as $\mathbf{x}=(\mathbf{BA})^{-1}\mathbf{B}\mathbf{y}(p)$, with $x_{m+1}=0$. Denoting a particular estimate by $\mathbf{x}^*$ and the matrix $(\mathbf{BA})^{-1}\mathbf{B}$ by $\mathbf{D}$, we immediately obtain $\mathbf{x}^*=\mathbf{D}(\mathbf{A}\mathbf{x}+\boldsymbol{\epsilon})=\mathbf{x}+\mathbf{D}\boldsymbol{\epsilon}$. By definition, the variance–covariance matrix of $\mathbf{x}^*$ is

$$\mathbf{V}(\mathbf{x}^*)=\mathrm{E}[(\mathbf{x}^*-\mathbf{x})(\mathbf{x}^*-\mathbf{x})^T],$$

i.e.

$$\mathbf{V}(\mathbf{x}^*)=\sigma^2\mathbf{D}\mathbf{D}^T.$$

In the least-squares case $\mathbf{B}=\mathbf{A}^T$, and $\mathbf{V}(\mathbf{x}^*)$ reduces to the familiar expression $\mathbf{V}(\mathbf{x}^*)=\sigma^2(\mathbf{A}^T\mathbf{A})^{-1}$. Since x_{m+1} is forcibly set to zero, an uncertainty Δx^*_{m+1} in x^*_{m+1} corresponds directly to a change $-\Delta x^*_{m+1}$ in the constant term in the Taylor expansion of $x_{m+1}(p)$. Such a change in the constant term is related to a change in p by an amount Δp from the computed value p_0, given as

$$\Delta p \approx \frac{\Delta x^*_{m+1}}{x'_{m+1}(p_0)}.$$

IV. APPLICATIONS

In this chapter we will discuss some applications and limitations of the proposed algorithm.

(1) When using the program given in Fig. 6 on the example in Chapter II, which as we saw proved to be a rather difficult task for the customary methods, no difficulties were encountered at all, and the program gave a very fast and accurate determination of the lowest minimum. If, for instance, a ninth order Taylor expansion around $p=0$ was performed and the function s^2 was used to form the artificial dimension, the value obtained for the nonlinear parameter differed only by about 0.1 % from the exact value of the lowest minimum, a difference which of course is far within the uncertainty of the order of 5 % in p inherent in the problem due to the statistical scatter in the input data. To illustrate the statement in Chapter III that there is a multitude of possible functions available for the construction of the artificial dimension, parameter estimates were also made using the functions s^l, $1/s^l$, $\sqrt[l]{s}$, $\cos^l s$, $\ln^l s$, with $l=3$, 4, 5, 6, 7 to form the artificial dimension; in all these cases the value obtained for the nonlinear parameter was far within the statistical fluctuations in p just as in the case of s^2 above.

(2) The algorithm assumes, in the formulation of Chapter III, the functions $\mathbf{y}(p)$ to be given as Taylor expansions of arbitrary order in p, i.e. $\mathbf{y}(p)=\mathbf{y}_0+\mathbf{y}_1p+\mathbf{y}_2p^2+\ldots$. In some cases an expansion of the type

$$\mathbf{y}(p)=\frac{\mathbf{y}_0}{p^l}+\frac{\mathbf{y}_1}{p^{l-1}}+\ldots+\mathbf{y}_l+\mathbf{y}_{l+1}p+\mathbf{y}_{l+2}p^2+\ldots$$

may be more suitable to approximate the functional behavior of $\mathbf{y}(p)$. This case is easily reduced to the one previously discussed by multiplying both sides in the approximation problem, Eqn. (1), by the unknown constant p^l (since p^l is constant, this merely corresponds to a change of all weights by the same amount), after which the algorithm can be used to calculate $p^l\mathbf{x}$ and p.

(3) An important limitation to the applicability of the algorithm in the form presented here is the requirement that the problem can be brought on the form $\mathbf{Ax}=\mathbf{y}(p)+\boldsymbol{\epsilon}$, i.e. to separate the linear and nonlinear parameters. This means that an approximation problem of the type, say, $f(s)\approx a+bs+ce^{\alpha s}+de^{\alpha s^2}$, where we want to determine a, b, c, d, and α, cannot be solved using the algorithm described here. (A problem of the type $g(s)\approx a+bs+ce^{\alpha s}$ is, however, after division by c easily reduced to a form $g(s)/c-a/c-bs/c\approx e^{\alpha s}$, for which the algorithm is applicable.)

(4) Potential limitations to the applicability of the algorithm are also inherent in the construction of the artificial dimension as discussed in Chapter III and the convergence properties of the power series expansion of the right-hand side. For reasons of numerical accuracy alone, one aims in general to have a number of observations which is large compared to the number of parameters to be determined; to find a function which does not improve the approximation will then as a rule, and as we also saw in paragraph (1) above, offer no problem. Also the convergence properties of the series expansion is in practice seldom any serious limitation, since the algorithm uses the projection $\mathbf{c}(p)$ of $\mathbf{y}(p)$ on $\mathcal{H}'_A$ rather than $\mathbf{y}(p)$ itself, and the projection has in general a more tranquil behavior than the original function.

V. CONCLUDING REMARKS

The noniterative algorithm presented in this paper might, if used within the limitations discussed in Chapter IV, be a useful alternative to the customary methods for solving approximation problems containing one nonlinear parameter.

In a general problem with several linear and nonlinear parameters

$$\mathbf{A}(\mathbf{p})\mathbf{x}=\mathbf{y}(\mathbf{p})+\boldsymbol{\epsilon},$$
$$\mathbf{B}(\mathbf{p})\boldsymbol{\epsilon}=0,$$

the solution of the linear parameters $\mathbf{x}$ should always in virtue of their uniqueness be separated from the solution of the nonlinear parameters $\mathbf{p}$. Especially simple is this in the case of gradient-free algorithms such as the one given by Powell (*loc. cit.*). In this case the criterion quantity

$$Q(\mathbf{x},\mathbf{p})=[\mathbf{A}(\mathbf{p})\mathbf{x}-\mathbf{y}(\mathbf{p})]^T\mathbf{M}[\mathbf{A}(\mathbf{p})\mathbf{x}-\mathbf{y}(\mathbf{p})]$$

can be reformulated as

$$Q(\mathbf{p})=[\mathbf{A}(\mathbf{p})\mathbf{x}^*(\mathbf{p})-\mathbf{y}(\mathbf{p})]^T\mathbf{M}[\mathbf{A}(\mathbf{p})\mathbf{x}^*(\mathbf{p})-\mathbf{y}(\mathbf{p})],$$

where the solution of the linear parameters,

$$\mathbf{x}^*(\mathbf{p})=[\mathbf{B}(\mathbf{p})\mathbf{A}(\mathbf{p})]^{-1}\mathbf{B}(\mathbf{p})\mathbf{y}(\mathbf{p}),$$

is calculated in the function-evaluating routine in the program. Here the present algorithm might be introduced instead to calculate both $\mathbf{x}^*$ and one of the nonlinear parameters, thus reducing the dimensionality of the nonlinearity, which may lead to a substantial increase in safety as well as computational economy.

ACKNOWLEDGEMENT

The author wishes to express his gratitude to Dr H. O. Zetterström for many clarifying discussions and to Prof. Åke Björk for several valuable suggestions to improvements of the presentation.

REFERENCES

Davidon, W. C., *Variable metric method for minimization.* A.E.C. Research and Development Report ANL-5990 Rev. (1959).

Faddeeva, V. N., *Computational Methods of Linear Algebra*, § 15. Dover, New York 1959.

Golub, G., Numerical methods for solving linear least-squares problems. *Numer. Math.* 7, 206–216 (1965).

Householder, A., *Principles of Numerical Analysis*, Ch. 6. McGraw-Hill, New York 1953.

IBM, *System/360 Scientific Subroutine Package*, Version III. IBM Manual GH20-0205-4 (1970).

Powell, M. J. D., A method for minimizing a sum of squares of non-linear functions without calculating derivatives. *Computer J.* 7, 303–307 (1965).

FOA REPORTS

Vols. 1 (1967), 2 (1968), 3 (1969), 4 (1970), and 5 (1971), see back cover of earlier issues.

Vol. 6 (1972)

No. 1 Evald von Zeipel, Axially Symmetric Air-Core Coils with Homogeneous and Strong Magnetic Field (FOA 2 Rapport B 2031-E1)

No. 2 Peter U. Tamm, Design of Multistage Transistor Amplifiers Using Scattering Parameters (FOA 3 Rapport B 3028-E1)

No. 3 Peter Erman, Gordon Berry, Lennart Lundin and Bertil Sigfridsson, Vacancy Cascade Spectroscopy: A Convenient Way of Studying Spectra and Lifetimes of Highly Ionized Atoms and Molecules (FOA 4 Rapport B 4053-A3)

No. 4 Lars-Henrik Andersson and Elise Halloff, Spectrophotometric Determination of Iron in Plutonium with Bathophenanthroline (FOA 4 Rapport B 4054-A29)

No. 5 Per Atterby, Plastics on Fire—Corrosion (FOA 1 Rapport B 1202-G1)

No. 6 Sylve Arnzén, An Experimental Investigation of the Accuracy of the Omega System in Sweden (FOA 3 Rapport B 3029-E4)

No. 7 Gunnar Walinder, Beta Doses in Wet Tissue (FOA 1 Rapport B 1205-A3)

No. 8 Bruno Slettenmark, Superconducting Quantum Interference Devices. I. Promising Designs for Stable Partly Resistive Quantum Interference Devices and an Easily Constructed Superconducting Quantum Interference Device (FOA 2 Rapport B 2034-E1)

No. 9 J. Gavin Park and Bruno Slettenmark, Superconducting Quantum Interference Devices. II. A Model for the Operation of a Partly Resistive Superconducting Point Contact Quantum Interference Device (FOA 2 Rapport B 2035-E1)

No. 10 Knut-Fredrik Alm and Lars-Henrik Andersson, Determination of Hydrogen, Carbon and Nitrogen in Plutonium Metal by a Combustion Technique and of Hydrogen by a Hot-Extraction Technique (FOA 4 Rapport B 4057-A2)

No. 11 Carl-Olov Criborn, Correlation Between Man and Mouse in Respect of Physical Activity and Oxygen Consumption (FOA 1 Rapport B 1233-A5(B2,C2))

No. 12 Elise Halloff, Spectrophotometric Investigation of Plutonium(III) Ions in Aqueous Solutions of Various Acids (FOA 4 Rapport B 4058-A2)

No. 13 Anita Meyerhöffer, The Molecular Structure of Some Anticholinergic Drugs (FOA 1 Rapport B 1242-C2)

Vol. 7 (1973)

No. 1 Johan Lundin, Considerations on a Chemical Arms Control Treaty and the Concept of Amplified Verification (FOA 1 Rapport B 1256-C1)

No. 2 Per Atterby and Lars Schön, An Electrochemical Method for the Investigation of Pitting on Mild Steel (FOA 1 Rapport B 1243-G1(G3))

No. 3 Arne Bergström, Noniterative Estimation of a Nonlinear Parameter (FOA 4 Rapport B 4059-M4)

Kr 3:–

Cover printed by FOA Repro
Stockholm 1973
Printed in Sweden

ISBN-91-7056-023-4