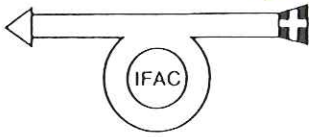
85-05C

PREPRINTS



3rd IFAC/IFIP INTERNATIONAL SYMPOSIUM

CADCE '85

Computer Aided Design in Control and Engineering Systems

ADVANCED TOOLS FOR MODERN TECHNOLOGY

Editors:

N.E. HANSEN and P. MARTIN LARSEN

THE TECHNICAL UNIVERSITY OF DENMARK
JULY 31 — AUGUST 2, 1985
LYNGBY COPENHAGEN DENMARK

A NUMERICALLY RELIABLE ALGORITHM FOR FITTING A SUM OF EXPONENTIALS OR SINUSOIDS TO NOISY DATA [xx]

B. De Moor[x], J. Vandewalle

ESAT Laboratory, Department of Electrical Engineering, Katholieke Universiteit Leuven
Heverlee, Belgium

Abstract. Many problems of practical CAD importance in systems, modelling and control can be formulated mathematically as a problem in signal decomposition where the signal is modelled as a finite number of independent elementary signals which are additively combined and corrupted by an additive zero mean white noise source. In this paper the parameters of a sequence of real valued and equidistant data of the form

$$y_k = \sum_{i=1}^{n} a_i e^{\beta_i k} e^{j(\gamma_i k + \delta_i)} + \eta_k \qquad k=0,N$$

where $\{\eta_k\}_{0,N}$ is a white zero mean noise sequence, are estimated using a realization technique, based upon the singular value decomposition of certain matrices. First, several possible practical combinations of the parameters α_i, β_i, γ_i and δ_i are considered with respect to their practical application. Secondly, the signal is modelled using a time-invariant autonomous state space model. Furthermore, some theoretical results are established which lead to a state space realization procedure that estimates the unknown parameters in a numerically reliable way. When the data are noise corrupted, it is shown that the techniques used still deliver remarkably reliable results, that however can be shown to be nearly optimal in a least squares sense. Furthermore, some numerical simulations reveal the reliability and possible interactive development of the algorithm.

Keywords. Computational methods, Harmonic analysis, Numerical methods, Parameter estimation, Signal processing, Singular value decomposition, Spectral analysis, State space method, Sum of exponentials.

INTRODUCTION

Let's consider a sequence of equidistant, real valued data, generated by the following expression

$$y_k = \sum_{i=1}^{n} \alpha_i e^{\beta_i k} e^{j(\gamma_i k + \delta_i)} \qquad (1)$$

$$k = 0, N$$

where $\alpha_i, \beta_i, \gamma_i, \delta_i \in \mathbb{R}$

The n terms of the sum can be considered as "elementary signals", each of them is characterized by four parameters : the amplitude α_i, a damping factor β_i, a pulsation γ_i that corresponds with an oscillation frequency $\delta_i = \gamma_i/2\pi$ and a phase δ_i (in radians).

Then the following special cases can occur

I. $\gamma_i = \delta_i = 0$ for $i = 1,\ldots,n$ then $y_k = \sum_{i=1}^{n} \alpha_i e^{\beta_i k}$ (2) is a sum of n real exponentials.

II. $\gamma_i \neq 0$ for $i = 1, n$; the n is even for real data.

Since the data y_k are assumed to be real, the parameters $\alpha_i, \beta_i, \gamma_i, \delta_i$, must occur in conjugated pairs, i. e. for each quadruplet (α_i, β_i, γ_i, δ_i,) there is a corresponding one $(\alpha_i, \beta_i, -\gamma_i, -\delta_i)$

For $l = 1 .. n/2$ we put

$$\begin{aligned} a_l &= 2\alpha_{2l} = 2\alpha_{2l-1} \\ b_l &= \beta_{2l} = \beta_{2l-1} \\ \omega_l &= \gamma_{2l} = \gamma_{2l-1} \\ \psi_l &= \delta_{2l} = \delta_{2l-1} \end{aligned} \qquad (3)$$

Then using Euler's formula the summation (1) is written as a sum of cosinoids with varying amplitudes.

$$y_k = \sum_{l=1}^{n/2} a_l e^{b_l k} \cos(\omega_l k + \psi_l) \qquad (4)$$

When $b_l < 0$, the corresponding elementary signal is damped. When $b_l = 0$, the elementary signal is a pure cosine.

III. Not all $\gamma_i \neq 0$ for $i = 1,n$ i.e. there are some i for which $\gamma_i = \delta_i = 0$. The general signal (1) now becomes a mixed sum of real exponentials and cosines.

A compact formulation of the problem is the following .

* Sponsored by a grant from the Belgian IWONL.

xx This work was partially supported as a project of the Belgian FKFO.

For a given n, minimize over all $\alpha_i, \beta_i, \gamma_i, \delta_i \in IR$

$$\sum_{k=0}^{N} w_k \left[\left(\sum_{i=1}^{n} \alpha_i e^{\beta_i k} e^{j(\gamma_i k + \delta_i)} \right) - y_k \right]^2 \quad (5)$$

for a given set of N +1 real valued and equidistant y_k and for a known set of weights w_k. The determination of the optimal "n" is considered furtheron. Signals of the general form (1) occur frequently in various applications. In fact, the general solution to any phenomenon described by a system of linear differential equations with constant coefficients is of the form (1) .

For the discussion of the statitistical question, whether it is justified to approximate a given data series by an exponential sum of the form (1), we refer to the literature [Ruhe, 1980] .

Although this contribution presents a unifying approach to the several 'special cases' I, II, III, generally, they are treated separately. Here we also stress the numerical aspects and the noise rejection qualities of the algorithm . The first solution to the problem of fitting experimental data by a sum of real exponentials is due to de Prony [1795] . Variations on his method are described in several books on numerical analysis [Lanczos, 1957] ('problem of "weighted moments" '). The problem occurs in fitting radioactive decay measurements (see ref. in [Ruhe, 1980]), transmission function analysis in atmospheric research [Wiscombe, 1977] compartmental analysis [Jacquez, 1972], electrical network analysis (Heaviside's expansion theorem [Lanczos, 1957], analysis of human lung exhalation and chemical rate constant estimation (see ref in [Ruhe, 1980] and [Wiscombe 1977]).

In many of these applications, the amplitudes a_i are expected to be positive. In [Wiscombe, 1977] a survey of various methods is given. The fundamental theorems guaranteeing existence and uniqueness of positive sums of a 'best fit' are treated by Cantor & Evans (1970). Algorithms based upon these results are described in [Wiscombe, 1977] and [Evans, Gragg, Leveque, 1980] . In [Ruhe, 1980] a non linear weighted least squares approach is studied while in [Ruhe, Wedin 1977] the problem is considered as a separable non-linear least squares problem since it is linear in the amplitudes a_i once the exponents β_i are known. A Remes-type algorithm is developped in [Braess, 1970] . while a Padé-approximant approach together with conditions for existence and uniqueness are discussed in [Sidi, 1982].

The case of fitting data by a sum of cosines is handled in [Lanczos, 1957] as the famous 'search for hidden periodicities'. However, more generally, the problem belongs to the domain of spectral analysis [Kay, Marple, 1981], where so called 'non traditional approaches' can perform much better then the Fast Fourier transforms in high resolution spectrum estimation with application to beamforming, direction finding, array signal processing and other areas. An excellent survey of various spectral estimation methods together with some 300 references, is given in [Kay, Marple, 1981]. The singular value decomposition as a key tool is used in [Tufts, Kumaresan, 1982] to obtain least squares approximations of limited rank of the signal correlation matrix estimate. In [Kung, 1981] a Toeplitz approximation method is presented which inspired this work.

The case of exponentially damped sinusoïds is considered in [Kumaresan, Tufts, 1982] where also the singular value decomposition of a matrix shows up.

Special care has to be taken with respect to the sometimes ill-conditioned nature of the 'sum of real exponentials'-problem. The parameters that describe the signal, are badly determined by the data. If these are perturbed by a small amount, there may be quite a large charge in the best-fit parameters. In [Lanczos, 1957] it is shown how several different exponential sums can approximate the same data series equally well. The fact that large variations are produced in the results by small variations in the data, is found in all methods and is an intrinsic property of the problem itself. It is customary to blame this ill-conditioning on the non orthogonality of the set of exponential functions [Lanczos, 1957]. However, a detailed sensivity analysis in [Ruhe, 1980] and [Wiscombe, 1977] illuminates the nature of the ill-conditioning more precisely.

This work fits in a general trend towards numerically reliable software for systems and control.

STATE SPACE MODEL AND SOME USEFUL PROPERTIES.

The problem of determining the parameters of the signal (4) , can be reformulated mathematically as a linear system realization problem. The signal of the general form (1) can be modelled by an autonomous linear state space model of dimension n

$$\begin{aligned} x_{k+1} &= A \cdot x_k \\ y_k &= C \cdot x_k \end{aligned} \quad (6)$$

Where

$$A = \begin{bmatrix} e^{\beta_1 + j\gamma_1} & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & e^{\beta_n + j\gamma_n} \end{bmatrix} \quad (7)$$

$$C = [\alpha_1 e^{j\delta_1} \quad \cdots \quad \alpha_n e^{j\delta_n}] \quad (8)$$

$$x_o = [1 \; 1 \; \ldots \; 1]^t \quad (9)$$

Proof :obvious by repeated substitution.

The key observation to understand the solution presented further, is to recognize that the general signal (1) is the homogeneous solution to a constant coefficient linear difference equation,

$$y_k = \sum_{i=1}^{n} a_i y_{k-i} , \qquad k = n, N \quad (10)$$

For a derivation of this expression we refer to [Kay, Marple, 1981]

The coefficients a_i are related to the elementary signals by Vieta's polynomial root theorem. From (10), it directly follows that the pxq Hankelmatrix with p, q> n, consisting of the p + q - 1 first data y_k, must have a rank equal to n.

$$H_{pq} = \begin{bmatrix} y_o & y_1 & y_2 & \cdots & y_{q-1} \\ y_1 & y_2 & \cdots & y_q & \\ y_2 & \cdots & & & \\ \cdot & & & & \\ \cdot & & & & \\ \cdot & & & & \\ y_{p-1} & \cdots & & & y_{p+q-2} \end{bmatrix} \quad (11)$$

and

$$\text{rank } [H_{pq}] = n \quad (12)$$

Proof : Any m-th column (m > n) can be written as a linear combination of the n preceding columns using (10) .

Further more, from the state space model (6) it follows that

$$y_k = C \cdot A^{k-1} \cdot X_0 \quad (13)$$

This leads to a decomposition of the Hankelmatrix H_{pq} :

$$H_{pq} = \begin{bmatrix} y_0 \cdots y_{q-1} \\ \vdots \\ y_{p-1} \cdots y_{p+q-2} \end{bmatrix} = \begin{bmatrix} Cx_0 & CAx_0 & \cdots & CA^{q-1}x_0 \\ CAx_0 & \cdots & & \\ CA^{p-1}x_0 & \cdots & & CA^{p+q-2}x_0 \end{bmatrix}$$

$$= \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{p-1} \end{bmatrix} [\, x_0 \; Ax_0 \cdots A^{q-1}x_0 \,] = \Gamma \cdot \Delta \qquad (14)$$

defining the matrices Γ(pxn) and Δ(nxp) that are both obviously of rank n.

$$\text{rank}\,(H_{pq}) = \text{rank}\,\Gamma = \text{rank}\,\Delta = n \qquad (15)$$

Now a new Hankelmatrix $\vec{H}_{pq}$ is defined which is derived from the previous one by shifting it over one colum to the right, while inserting one new date sample y_{p+q-1}

$$\vec{H}_{pq} = \begin{bmatrix} y_1 & y_2 & \cdots & y_q \\ y_2 & \ddots & & \\ \vdots & & & \\ y_p & \cdots & & y_{p+q-1} \end{bmatrix} \qquad (16)$$

where again due to (10)

$$\text{rank}\,(\vec{H}_{pq}) = n \qquad (17)$$

Using (12) and (13), it's easy to prove that

$$\vec{H}_{pq} = \Gamma \cdot A \cdot \Delta \qquad (18)$$

and the least squares estimate of A is computed as

$$A = \Gamma^{+} \cdot \vec{H}_{pq} \cdot \Delta^{+} \qquad (19)$$

where Γ^{+} and Δ^{+} denote the pseudo-inverse .

DERIVATION OF THE ALGORITHM.

Following [Zeiger, Mc Ewen, 1974] and [Kung, 1978] the singular valuedecomposition of the data Hankelmatrix H_{pq} is used [Klema, Laub, 1980]

-to estimate the (numerical) rank and then the number of elementary signals to model.
-to obtain numerically reliable estimation of the matrices Γ, Δ, A and C.

Suppose the (partial) SVD of H_{pq} is given by

$$H_{pq} = \underset{pxn}{U} \cdot \underset{nxn}{\Sigma} \cdot \underset{nxq}{V^t} \qquad (20)$$

then the 'balanced realization technique' computes Γ and Δ up to an nxn similarity transform T, which is to be determined furtheron.

$$\Gamma \cdot T = U \cdot \Sigma^{1/2} \qquad (21)$$

and

$$T^{-1} \cdot \Delta = \Sigma^{1/2} \cdot V^t \qquad (22)$$

The least-squares estimate of A, up to the similarity transformation T, is computed using (19) as

$$T^{-1} \cdot A \cdot T = (\Gamma \cdot T)^{+} \cdot \vec{H}_{pq} \cdot (T^{-1} \cdot \Delta)^{+} = \Sigma^{-1/2} \cdot U^t \cdot \vec{H}_{pq} \cdot V \cdot \Sigma^{-1/2} \qquad (23)$$

Now the similarity transformation matrix T is computed by taking into consideration that

1) the systemmatrix A of the state space model(6) is diagonal and
2) the first column of the matrix Δ (14) is given by (9)

Since A and $T^{-1} \cdot A \cdot T$ are similar, they share the same eigenvalues. So condition 1) is fulfilled by computing the eigenvalues and -vectors of $T^{-1} \cdot A \cdot T$.

Let X be the nxn matrix containing as columvectors the eigenvectors of $T^{-1} \cdot A \cdot T$. and Λ be the diagonal nxn matrix with the eigenvalues, then clearly

$$T^{-1} \cdot A \cdot T = X \cdot \Lambda \cdot X^{-1} \qquad (24)$$

and using (18)

$$\vec{H}_{pq} = (\Gamma \cdot T) \cdot (T^{-1} \cdot A \cdot T) \cdot (T^{-1} \cdot \Delta) = U \cdot \Sigma^{1/2} \cdot X \cdot \Lambda \cdot X^{-1} \cdot \Sigma^{1/2} \cdot V^t = (U \cdot \Sigma^{1/2} \cdot X) \cdot \Lambda \cdot (X^{-1} \cdot \Sigma^{1/2} \cdot V^t) \qquad (25)$$

Now let c denote the (complex) first column of $X^{-1} \cdot \Sigma^{1/2} \cdot V^t$

$$c = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = X^{-1} \cdot \Sigma^{1/2} \cdot V^t \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \qquad (26)$$

Define by D^{-1} the complex diagonal nxn matrix containing the inverse entries of c

$$D^{-1} = \begin{bmatrix} \frac{1}{c_1} & 0 & \cdots & 0 \\ \vdots & \ddots & & \vdots \\ 0 & \cdots & & \frac{1}{c_n} \end{bmatrix} \qquad (27)$$

Then it is obvious that

$$D^{-1} \cdot c = x_0 = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \\ 1 \end{bmatrix} \qquad (28)$$

Hence

$$D^{-1} \cdot X^{-1} \cdot \Sigma^{1/2} \cdot V^t \cdot \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \qquad (29)$$

and (24) is to be modified as

$$\vec{H}_{pq} = (U \cdot \Sigma^{1/2} \cdot X \cdot D) \cdot \Lambda \cdot (D^{-1} \cdot X^{-1} \cdot \Sigma^{1/2} \cdot V^t) \qquad (30)$$

which finally results in the parameter estimation

$$A = \Lambda = \begin{bmatrix} e^{\beta_1 + j\gamma_1} & & \\ & \ddots & \\ & & e^{\beta_n + j\gamma_n} \end{bmatrix} \qquad (31)$$

and C as the first row of $U \cdot \Sigma^{1/2} \cdot X \cdot D$

$$C = [1\ 0 \ldots 0] \cdot U \cdot \Sigma^{1/2} \cdot X \cdot D = [\alpha_1 e^{j\delta_1} \ldots \alpha_n e^{j\delta_n}] \qquad (32)$$

In conclusion of this section we summarize the algorithm

Step 1 : Construct the matrices H_{pq} and $\vec{H}_{pq}$ (eq. 11,16)

Step 2 : Compute SVD of H_{pq} (eq. 20)

Step 3 : Compute X.Λ from (23), (24) via an eigenvalue decomposition.

Step 4 : Determine all parameters from (31), (32).

COMPUTATIONAL CONSIDERATIONS

We now consider some computational refinements. The singular value decomposition (20) and the complex eigenvalue decomposition (24) can be computed using fully tested standard subroutines.

In the special case of real exponentials and when square Hankelmatrices are used (p=q), the realized matrix $T^{-1} \cdot A \cdot T$ can be proven to be symmetric and definite positive, making it's eigenvalue decomposition well condicioned [Wilkinson, 1965]

In the other cases and using square Hankelmatrices, the realized matrix $T^{-1}.A.T$ is symmetric when the absolute value of its entries are considered.
The diagonal matrix (27) is computed as the complex solution to the set of equations defined by (29)

$$X.t = \Sigma^{1/2}.V^t . \begin{bmatrix} 1 \\ 0 \\ . \\ . \\ . \\ 0 \end{bmatrix} \qquad (33)$$

Equating real and imaginary parts of (33)

$$X = X_r + jX_i$$

$$t = t_r + jt_i$$

$$\begin{bmatrix} X_r & -X_i \\ X_i & X_r \end{bmatrix} . \begin{bmatrix} t_r \\ t_i \end{bmatrix} = \Sigma^{1/2}.V.t. \begin{bmatrix} 1 \\ 0 \\ . \\ 0 \end{bmatrix} \qquad (34)$$

This set can be solved by another SVD. Despite the additional computational requirements, the condition of (34) can be monitored.
Having calculated the matrices A (31) and C (32) the parameters α_i, β_i, γ_i, δ_i, are computed as follows

$$A = \text{diag } (\lambda_i) \text{ then } \beta_i = |\lambda_i| \qquad (35)$$

$$\delta_i = \arg (\lambda_i)$$

$$C = [\alpha_1 e^{j\delta_1} \ldots \alpha_n e^{j\delta_n}] \text{ then } \alpha_i = |\alpha_i e^{j\delta_i}|$$

$$\delta_i = \arg (\alpha_i e^{j\delta_i}) \qquad (36)$$

INFLUENCE OF THE NOISE : SUBOPTIMALITY

So far, we only considered the case in which
I. the data were known to be exactly representable by a finite sum of elementary signals (1)
and II. no noise corrupted the data.
However, when the data are additively corrupted by noise with variance σ^2, several effects that are difficult to separate, cause the method presented here to be suboptimal in the sense defined by (5).
-If n elementary signals are involved causing the noiseless Hankelmatrix H_{pq} to have n singular values to be non-equal to zero

$$\sigma_1 \geq \sigma_2 \geq \ldots . \geq \sigma_n > 0$$

and (min (p,q) - n) to equal zero, the noise corrupted version of this Hankelmatrix will have <u>all</u> its singular values σ_i different from zero. It is proven in [Anderson, 1963] that, asymptotically, as $N \to \infty$, the eigenvalues of a perturbed covariance matrix behave like

$$\text{for } i = 1, n \quad \tilde{\lambda}_i = \lambda_i + \sigma^2$$
$$\text{and for } i = n+1, \min (p,q) \quad \tilde{\lambda}_i = \sigma^2 \qquad (37)$$

and also some likelihood ratio criterions are proposed in order to obtain the best rank estimation.
A similar approach is followed in [Kung, 1981] where, due to the same criterion (37) the ambient space is divided into a n-dimensional signal subspace and a [min (p,q) - n] - dimensional noise subspace. Another reliable rank criterium is derived [De Moor, 1984] and [Staar, 1982]. Only those singular values are retained that satisfy

$$\tilde{\sigma_i} > \sqrt{(R_o+1) \max (p,q)} \cdot \sigma \qquad (38)$$

where R_o is a user defined maximal minimal signal to noise ratio (SNR).
The corresponding singular vectors form an estimate of the signal subspace.
It is obvious that if one of the signal components falls below the level of accuracy of the date points it is completely unrecoverable. This is an ever present caveat in the problem of separating elementary signals.
-If the numerical rank n is estimated, using one of the preceding criterions, the noise corrupted Hankelmatrix of full rank is approximated in a least squares sense by a matrix computed from the partial SVD (20), according to the theory developped in [Eckart, Young, 1936]. However, this matrix generically shows no longer the Hankelstructure and hence does not represent a n-th order linear system (6). Although this problem is not yet solved for the least squares case with Frobenius-norm, the approximation, obtained using the method here presented, can be shown to be close to optimal [Kung, 1978,1981]. This is due to the well known insensitivity of the singular vector subspaces to perturbations in the data [Wilkinson, 1965].
-Although the singular value decomposition is well conditioned , the singular vectors and values differ from the original ones due to the noise. Usually, the effect on the singular vectors is neglected, not only because of the insensitivity of the signal subspace but also because the effect of the perturbation is rather complicated, as is shown by a first order perturbation analysis of the SVD in [De Moor, 1984].
Concerning the perturbations of the singular values practical tests have shown that an improvement in the results (reduction in bias) is to be expected, when the partial SVD (20) is replaced by

$$\hat{H}_{pq} = U.[\Sigma - S].V^t \qquad (39)$$

where S is a diagonal matrix.

$$S = \text{diag } (\hat{\sigma})$$

Here $\hat{\sigma}$ is an estimate of the supposed isotropic noise singular values.
(maximum likehood estimate [Anderson, 1963], arithmetic mean [Kumaresan, Tufts, 1982]).
We conclude that the noise causes the algorithm to be suboptimal but on the other hand, the partial SVD increases the signal to noise ratio dramatically because the date are projected into the signal subspace. The noise compensation procedure produces a reduction in bias of the estimates.

NUMERICAL RESULTS

In this section we present some striking numerical results that illustrate the power of the algorithm.
I. Fitting a sum of exponentials.
In [Wiscombe, Evans, 1977] it is reported that all known methods experience great difficulty in reconstructing a particular exponential sum, given data sampled from that sum. The ability of our algorithm is however excellent in doing this job, even if the data are only accurate to 2 decimal digits (rounded). 97 data points of the sum

$$y_k = 0.1\, e^{-0.01k} + 0.3\, e^{-0.1k} + 0.6\, e^{-k}$$
$$= a_1.b_1^k + a_2.b_2^k + a_3\, b_3^k$$

were generated for k=0,96 and stored in 48x48 Hankelmatrices.
5runs with the algorithm using unrounded data and data rounded to 5,4,3 and 2 decimal digits were performed on an IBM 3033 machine with 24 bit mantissa and 7 bit exponent. Following quantities were computed : coefficients and exponents $\hat{x}$, relative deviation defined as 100. $(\hat{x}_{estimate} - x_{exact}) / x_{exact}$, and the error measures

$$e_{max} = \max_{0<k<96} | y_k - \hat{y}_k |$$

$$e_{rms} = \left[\sum_{k=1}^{96} [y_k - \hat{y}_k]^2 \right]^{1/2} \Big/ 96$$

$$e_{\%} = (100/96) \sum_{i=1}^{96} |y_k - \hat{y}_k| / |y_k|$$

where $\hat{y}_k$ denotes the estimate of the k-th value of the sum.
The results are listed in table 1-3

TABLE 1 ACCURACY OF ESTIMATED PARAMETERS IN %

	6D	5D	4D	3D	2D
a_1	-0.01	-0.02	-0.3	-0.4	0.6
b_1	-0.0001	0.0003	0.0004	-0.001	-0.008
a_2	0.006	0.01	0.06	-0.09	0.1
b_2	0.0004	0.0006	0.002	0.01	-0.07
a_3	-0.0006	0.001	0.002	0.07	-0.1
b_3	-0.0006	-0.0002	0.0003	0.1	-0.3

TABLE 2 ERROR MEASURES

	6D	5D	4D	3D	2D
e_{rms}	0.21E-06	0.23E-06	0.47E-06	0.39E-05	0.35E-04
$e_{\%}$	0.24E-02	0.33E-02	0.41E-02	0.12E-01	0.95E-01
e_{max}	0.36E-05	0.38E-05	0.12E-04	0.17E-03	0.26E-02

TABLE 3 SINGULAR VALUES OF THE 48x48 HANKELMATRICES

	6D	5D	4D	3D	2D
σ_1	4.319	4.319	4.319	4.319	4.315
σ_2	0.890	0.890	0.890	0.890	0.889
σ_3	0.256	0.256	0.256	0.256	0.257
σ_4	$\sim 10^{-5}$	$\sim 10^{-4}$	$\sim 10^{-4}$	$\sim 10^{-3}$	$\sim 10^{-2}$

Comparing the results listed in table 1 and 2 to those reported in [Wiscombe, Evans, 1977], it can be stated that the algorithm we proposed, obtains the same level of accuracy. Table 3 confirms the fact that the singular values are quite insensitive to perturbation in the data.
It's interesting to note that the choice of the Hankelmatrix dimensions is quite arbitrary as long as the minimal dimension is larger than the number n of elementary signals. Several tests on matrices with large varying dimensions, could detect no significant difference in the accuracy of the estimate.

II. Sum of (damped) cosines.
In order to illustrate the reliability, 91 data were sampled from the four term damped cosine series

$$y_k = 21.(0.9)^k \cos(0.2k+0.7) + 7.(0.8)^k \cos(0.1k+0.4) + 38.(0.4)^k \cos(0.45k+0.8) + 20.(0.9)^k \cos(0.3k+0.9)$$
$$= \sum_{i=1}^{4} \alpha_i \beta_i^k \cos(\gamma_i k + \delta_i)$$

The data were stored in a 45X45 Hankelmatrix with singular spectrum

σ_1=99.68; σ_2=81.04; σ_3=9.881; σ_4=4.328; σ_5=3.377; σ_6=0.2933; σ_7=0.1941E-01; σ_8=0.1857E-01

and the remaining ones all smaller then 0.15E-03. This leads to the obvious decision that the Hankelmatrix rank must be 8.
Full precision accuracy results are listed in Table 4 and 5 .

TABLE 4 ACCURACY OF ESTIMATED PARAMETERS IN %

i	1	2	3	4
α_i	0.09	0.3	0.03	0.1
β_i	-0.003	-0.08	0.005	-0.003
γ_i	0.03	-0.2	0.04	-0.01
δ_i	-0.4	1.1	0.05	0.05

The error measures are :
e_{rms} = 0.208E-04
$e_{\%}$ = 0.397
e_{max} = 0.643E-03
These results are excellent, taking into consideration that machineprecision corresponds to 6 à 7 decimal digits.

III. Even the identification of so-called "unstable" signals causes no problem as long as overflow is avoided.
Sampled data from the sum :

$$y_k = 10.(1.1)^k.\cos(0.4k + 0.6) + 7.(0.9)^k \cos(0.2k + 0.4) = \sum_{i=1}^{2} \alpha_i \beta_i^k \cos(\gamma_i k + \delta_i)$$

were additively corrupted with a pseudorandom white, zero mean, Gaussian noise sequence with variance σ^2=1 .
51 data samples were used and the Hankelmatrix dimensions were 25x25.
The singular values are σ_1=3434.5; σ_2=2108.5; σ_3=18.511; σ_4=15.214; σ_5=6.8939; σ_6=6.1657; σ_7=...
From σ_5 on, the singular values reach a saturation level, indicating that the noise level is reached.
The accuracy of the estimates, taking a rank estimate of n=4, is listed in Table 5.

TABLE 5 ACCURACY OF ESTIMATED PARAMETERS IN %

i	1	2
α_i	0.005	0.6
β_i	-0.7	2.3
γ_i	0.005	-6.7
δ_i	-1.59	33.0

Remark that the errors for the second signal are larger than that of the first because the first signal contribution is larger than the second and this is less perturbed by the noise.
The error measures are :
e_{rms} = 0.814
$e_{\%}$ = 5.178
e_{max} =25.1
The last error being caused by the amplification of the misfitting due to the instability.

IV. In the following example the effect of additive white noise of increasing variance is investigated, using the sum

$$y_k = 10.(0.95)^k \cos(0.2k + 0.8) + 12(0.85)^k \cos(0.4k + 0.3)$$

41 data were generated and 20x20 Hankelmatrices were used.
The signal-to-noise ratio SNR is defines as

$$SNR = 10 \log \left[\left(\sum_{k=0}^{40} y_k^2 \right) / 41\sigma^2 \right] \quad [d3]$$

The error measures are listed in Table 6.

TABLE 6 ERRORS AS A FUNCTION OF THE NOISE LEVEL

	σ^2	SNR(dB)	e_{rms}	e_z	e_{max}
0	0	-	0.45E-05	0.35E-02	0.57E-04
1	0.0001	54	0.68E-03	0.51E+00	0.32E-02
2	0.0005	47	0.15E-02	1.14	0.183-01
3	0.001	44	0.22E-02	1.6	0.26E-01
4	0.005	37	0.47E-02	3.6	0.57E-01
5	0.01	34	0.67E-02	5.1	0.80E-01
6	0.05	27	0.15E-01	11.6	0.18
7	0.1	24	0.20E-01	16.7	0.26
8	0.5	17	0.49E-01	46.5	0.67
9	1.0	14	0.72E-01	71.2	0.98
10	2.0	11	0.84E-01	78.7	1.36
11	5.0	7	0.12	80.0	2.54

To conclude this section, we now mention some of the software aspects.
The program is written in FORTRAN. It's up to the user to choose between a fitting of the data I) by a sum of real exponentials II) by real cosines or III) a combination of the two. A list of the computed singular values is shown to let the user decide whether the numerical rank is to be determined automatically (using one of the rank criterions) or not. We want to stress here that this is the only decision step in the algorithm in contrast with e.g. convergence and coalesce decision to be made in the method of [Wiscombe, Evans, 1977]. At a simple request intermediate results are printed together with all kinds of a posteriori indicators of the numerical reliability (condition-numbers, scaling factors, first-order perturbation analysis using the estimated Jacobian etc..)

CONCLUSIONS

A numerically reliable algorithm for the estimation of the parameters of a complex sum of exponentials is derived. The data must be equidistant and may be corrupted by noise. The method is based upon the balaced realization technique of a linear state space model.
The key tool is the singular value decomposition. It guarantees a reliable estimation of the number of elementary signals and of the parameters.
Although the noise causes the algorithm to be suboptimal, some practical examples prove the method to have a more than sufficient degree of accuracy.
Further research will concentrate upon a detailed sensitivity analysis, extending the algorithm for complex valued data and taking into consideration a possible adaptive scheme based upon adaptive singular value decomposition techniques.

REFERENCES

Anderson, T.W. (1963). Asymptotic theory for principal component analysis. Ann. Math. Statist., 34, 122-148

Braess, D. (1970). Die Konstruktion der Tschebyscheff Approximierenden bei der Anpassung mit exponentialsummen. Journal of approximation theory, 3, 261-273

Cantor, D.G. and J.W. Evans (1970). On approximations by positive sums of powers. Siam J. Appl. Math., 18, 380.

De Moor, B. (1984). First order perturbation analysis of the singular value decomposition. Int. report, ESAT lab., K.U. Leuven

De Moor, B., J. Vandewalle and J. Vanderschoot (1984). A filtering principle for vector signals based on the singular value decomposition Proc. Europ. Conf. on Circuit Theory and Design (ECCTD), 1985, Prague. Submitted.

De Prony, R. (1795). Essay expérimentale et analytique. J. EC. Polytech. Paris I, 24-76

Dongarra, J.J., J.R. Bunch, C.B. Moler, and G.W. Stewart (1979). Linpack User's guide. SIAM Philadelphia.

Eckart, C. and G. Young (1936). The approximation of a matrix by another of lower rank. Psychometrika, vol 1, 211-218.

Evans, J.W, W.B. Gragg, and R.J. Leveque (1980) On least squares exponential sum approximation with positive coefficients. Math. Comp., 34, 203-212.

Jacquez, J.A. (1972). Compartmental analysis in biology and medecine. Elsevier Publishing Company, Amsterdam

Kay, S.M. and S.L. Marple (1981). Spectrum analysis, A modern perspective. Proc. IEEE, Vol. 69, no 11.

Klema, V.C. and A.J. Laub (1980). The singular value decomposition : its computation and some applications. IEEE Trans. Aut. Contr., Vol AC-25,2, 164-176.

Kumaresan, R. and D.W. Tufts (1982). Estimating the parameters of exponentially damped sinusoids and pole-zero modelling in noise. IEEE. Trans. on acoustics, speech and signal processing, Vol ASSP-30,6.

Kung, S.Y. (1978). A new identification and model reduction algorithm via singular value decomposition. Proc. 12th Asilomar Conf. Circuits, Syst. Computer., Pacific Grove, Calif., 705-714.

Kung, S.Y. (1981). A Toeplitz approximation method and some applications. Int. Symp. on math. theory of networks and systems, Vol. 4, Santa Monica, California.

Lanczos, C. (1956). Applied Analysis. Prentice Hall, Inc. Englewood Cliffs, N.J.

Ruhe, A. and Per Ake Wedin (1980a) Algorithms for separable nonlinear least squares problems SIAM review , Vol. 22, 3.

Sidi Avram (1982). Interpolation at equidistant points by a sum of exponential functions. Journ. of approximation theory, 34, 194-210

Staar, J. (1982). Concepts for reliable modelling of linear systems with application to online identification of multivariable state space descriptions. PhD thesis, ESAT Lab, K.U. Leuven.

Tufts, D.W. and R. Kumaresan (1982). Estimation of frequencies of multiple sinusoids: Making linear prediction perform like maximum likehood. Proc. IEEE, Vol. 70, 9.

Vandewalle, J., J. Staar, B. De Moor, J. Lauwers (1984). An adaptive singular value decomposition algorithm and its use in adaptive realization. 6th. int. conf. anal. optim. of systems, June 1984, Nice.

Wilkinson, J.H (1965). The algebraic eigenvalue problem. London, Oxford Univ. Press.

Wiscombe, W.J. and J.W. Evans (1977) Exponential sum fitting of radiative transmission functions. Journal of Comput. Physics, 24, 416-444.

Zeiger, H.P. and A.J. Mc Ewen (1974) Approximate linear realization of given dimension via Ho's algorithm. IEEE Trans. Aut. Control, Vol. AC-19, 153.