PARAMETRIC MODELING OF ELECTROMAGNETIC WAVEFORMS

Luc Knockaert
Laboratory of Electromagnetism and Acoustics, Ghent University
St. Pietersnieuwstraat 41, B-9000 Gent, Belgium.

Abstract

The problem of modeling and approximating a waveform by a linear combination of basis functions containing a variable parameter is considered. It is shown that the Kalman equation error concept of linear system identification theory can, in a modified form, be applied to a large class of modeling problems, provided the chosen basis function is a solution of a linear functional equation in Hilbert space. This class includes rational and Tauberian modeling problems, known to be of relevance for electromagnetic transient response and wide bandwidth radar return signature identification respectively.

1 INTRODUCTION

Modeling a waveform by a linear combination of basis functions containing a variable parameter, is not a new problem. Already by the end of the 18th century Gaspard Riche, Baron de Prony [1] was led to investigate the modeling of a function in terms of a linear combination of exponentials with different exponents, and proposed his celebrated method. One could say that this was the first contribution to the theory of rational modeling, since to a rational function with simple poles in the Laplace domain corresponds a weighted sum of exponentials in the time domain, representing the impulse response.

Interest in rational modeling schemes rose dramatically with the advent of the Singularity Expansion Method (SEM) [2] and Prony's method was in some sense rediscovered [3]. Of course, the modern versions of the technique, called 'extended Prony' [4] or 'semi least-squares' [5] method,

bear little resemblance to Prony's original approach. In addition, numerous other rational modeling techniques, iterative or non-iterative, in time or frequency domain, exist and have been extensively treated in the literature [6]-[10]. The final justification of the rational modeling approach is the fact that the SEM-poles or natural frequencies of a given scatterer are aspect and illumination independent, which makes them highly attractive features for target identification and discrimination purposes.

But not all waveforms can be interpreted as transients, and hence, depending on the physical content of any given problem, other than rational parametric models may be appropriate. For instance, in [11] a feature extraction procedure based on a Tauberian approximation is applied to a wide bandwidth radar return signal. By Tauberian approximation is meant a linear combination of shifted versions of a single known basis function. The rationale for using this kind of model in this particular case is the fact that the return waveform consists of a sequence of similar short pulses emanating from different scattering centers, areas of discontinuity, on the surface of the target. The target features to be extracted are clearly the relative delays and relative amplitudes of the pulses, with emphasis on the former, and this is implemented using a frequency domain Prony technique. The method proves quite successful for target discrimination and classification purposes. It may at first seem that rational and Tauberian modeling have nothing in common, but closer inspection reveals that the only — but essential — difference resides in the choice of the basis function and parameter. We have basically exponentials and exponents in the first case vs. pulses and delays in the second case.

The aim of this paper is to provide parametric modeling problems with more flexibility as to the choice of the basis functions. The basis function is not viewed as an isolated item but as a member of a category of functions defined in terms of a linear functional equation in Hilbert space, and the modeling procedure is based, not on the explicit form of the basis function, but on the implicitly defined form expressed by this functional equation. In order to carry out this programme, a formulation in terms of the Kalman equation error concept is necessary [12]-[13]. Section 2 formulates the modeling problem and introduces the error and Mean Squared Error (MSE) concept. Section 3 introduces the basic functional equation, the Kalman equation error and the Mean Squared Equation Error (MSEE) concept. In Section 4 the theory is applied to modeling by sinusoids and exponentials, rational and Tauberian approximation, and finally to modeling by Bessel functions.

2 PROBLEM FORMULATION : THE ERROR CONCEPT

Notation : in what follows we work in a Hilbert space $H(\Omega)$ of vectors (functions) with support Ω. The scalar product is (x, y) and the norm $= \|x\| = \sqrt{(x, x)}$. The adjoint of the linear operator K is denoted $K^\dagger$ and defined as usual, while $K(t, t')$ stands for the integral kernel of K. In a matrix context † means Hermitian conjugate.

A vector (waveform) v representing observed data is given, and considerations about the underlying physical process indicate that, in the absence of noise, v can be modeled as a sum of known basis vectors (functions) $y(s; c)$ depending non-linearly on the parameter s and linearly on the parameter c. In fact $y(s; c) = cy(s; 1)$, but by including the factor c explicitly, the term 'linear combination' can be replaced by 'sum'. This may seem artificial, but it somehow simplifies the notation.

In other words, there is a number n_1 and sets of parameters $\{c_1, \ldots, c_{n_1}\}$, $\{s_1, \ldots, s_{n_1}\}$ such that

$$v = \sum_{k=1}^{n_1} y(s_k; c_k) \tag{1}$$

and the problem is, given the data vector v, to extract the complex parameters c_k and s_k.
In order to have a unique representation of the form (1) we must require :
i) all c_k are non-vanishing
ii) all s_k are different
iii) linear independence i.e. if all s_k are different then
$\sum_{k=1}^{n} y(s_k; c_k) = 0$ implies that all c_k are vanishing.
Requirement iii is closely related to the concept of unisolvency [14]. The assumption of noiseless data is highly unrealistic and therefore a better formulation is :

$$v = \sum_{k=1}^{n} y(s_k; c_k) + \eta \tag{2}$$

with η representing the noise vector, which is supposed to be white, zero-mean and Gaussian with variance σ^2. The presence of noise changes the nature of the problem : while (1) represents an exact fitting, (2) is basically an estimation problem, and hence a solution based on Maximum Likelihood (ML) estimation theory can be formulated. The likelihood function for the white, zero-mean Gaussian noise vector η is proportional to $\exp(-\|\eta\|^2/2\sigma^2)$, and since

$$\eta = v - \sum_{k=1}^{n} y(s_k; c_k) \tag{3}$$

it is not difficult to see that ML estimation of the parameters s_k and c_k is equivalent to the following MSE minimization problem : find the parameters s_k and c_k that minimize the functional

$$\|v - \sum_{k=1}^{n} y(s_k; c_k)\|^2 \tag{4}$$

Since n_1 — referred to as the exact order — is not known in general, it seems preferable to let the order, say n to distinguish it from n_1, free, and follow a MSE approximation procedure for various n. Defining the error as

$$e_n = v - \sum_{k=1}^{n} y(s_k; c_k) \tag{5}$$

the functional to be minimized is

$$\|e_n\|^2 = \|v - \sum_{k=1}^{n} y(s_k; c_k)\|^2 \tag{6}$$

Note that the extracted parameters will depend on the particular choice of n. Furthermore it seems tempting to take as exact order the one corresponding to the minimum squared error, but this is wrong since it can be shown (see Appendix A) that the MSE sequence is ever-decreasing, i.e. $\|e_{n+1}\|^2 \leq \|e_n\|^2$ for all n. This shows that it is not easy to find a consistent order determination criterion, and therefore, in what follows, we will suppose that n is part of the data.

Finally, even with n fixed, the implementation of (6) still represents a formidable task, due to the presence of the non-linear parameters s_k. Of course this depends on the explicit functional form of $y(s; c)$, but even in the relatively simple exponential modeling case [15] ($y_t(s; c) = ce^{st}$,) the algorithms – which are necessarily iterative in nature — are plagued with convergence problems and yield local minima rather than the global one. This explains the need for an alternative formulation.

3 FUNCTIONAL EQUATION AND EQUATION ERROR

3.1 Functional equation. Definitions.

The fundamental hypothesis is that the basis vector $y(s; c)$ is a solution of a functional equation of the form

$$Ky = sy + cu \tag{7}$$

where K is a known linear operator and u is a known vector in $H(\Omega)$, with K and u independent of the parameters s, c. It is clear that y depends linearly on c and non-linearly on s.

The special case with $u \equiv 0$ will be treated separately. In that case we have $Ky = sy$.
Define the polynomials $p(z), q(z)$ and the linear operators $p(K), q(K)$ as

$$p(z) = p_n \prod_{k=1}^{n} (z - s_k) = \sum_{k=0}^{n} p_k z^k \tag{8}$$

$$q(z) = p_n \sum_{k=1}^{n} c_k \prod_{j \neq k} (z - s_j) = \sum_{k=0}^{n-1} q_k z^k \tag{9}$$

$$p(K) = \sum_{k=0}^{n} p_k K^k \tag{10}$$

$$q(K) = \sum_{k=0}^{n-1} q_k K^k \tag{11}$$

The leading coefficient p_n is supposed to be non-vanishing. The parameters s_k are the roots of $p(z)$, while the c_k are given by

$$c_k = q(s_k)/p'(s_k) \tag{12}$$

3.2 Main result. Equation error.

We are now able to state the main result :

$$p(K)e_n = p(K)v - q(K)u \tag{13}$$

Proof : Applying the operator

$$p(K) = \sum_{k=0}^{n} p_k K^k = p_n \prod_{k=1}^{n} (K - s_k I) \tag{14}$$

on both sides of (5) gives

$$p(K)e_n = p(K)v - p_n \prod_{k=1}^{n} (K - s_k I) \sum_{j=1}^{n} y(s_j; c_j) \tag{15}$$

Since all factors in the factorization (14) are commutative and since

$$(K - s_k I)y(s_k; c_k) = c_k u \tag{16}$$

by the fundamental hypothesis (7), it is clear that the formulas (15) and (13) are equivalent. □
We call the quantity

$$\tilde{e}_n = p(K)e_n \tag{17}$$

the equation error for the following reason. Suppose K represents the operator d/dt; then, after transforming into the Laplace domain, and assuming zero initial conditions, (13) is equivalent with

$$\tilde{E}_n(s) = p(s)V(s) - q(s)U(s) \tag{18}$$

— here s is the Laplace variable —, but now $\tilde{E}_n(s)$ is precisely the Kalman equation error [12]-[13], of a linear system with input U, output V and transfer function $H(s) = q(s)/p(s)$.

Hence, instead of minimizing the MSE, which is a complex non-linear problem, we tend to minimize the MSEE $\|\tilde{e}_n\|^2$ with respect to p_i and q_j, which is a linear problem. This can be seen as follows : let

$$v_k = K^k v, \qquad u_k = K^k u \tag{19}$$

then (13) can be reformulated as

$$\tilde{e}_n = \sum_{k=0}^{n} p_k v_k - \sum_{k=0}^{n-1} q_k u_k \tag{20}$$

Now since $\tilde{e}_n$ is linear in p_i, q_j, $\|\tilde{e}_n\|^2$ is quadratic in p_i, q_j, and it is well known that the minimization of a quadratic form results in linear equations. Once the p_i, q_j determined, the true parameters s_i, c_j are obtained by rooting $p(z)$ and substituting the roots s_i in (12). But one must be careful. In the case $u \equiv 0$ we have

$$\tilde{e}_n = \sum_{k=0}^{n} p_k v_k \tag{21}$$

and it is obvious that minimizing the MSEE in this case cannot possibly yield information about the parameters q_i or c_i.

Before proceeding, it is therefore necessary to split up the problem in two parts. A first part with $u \equiv 0$, called denominator estimation, and a second part with the set $\{u_0, u_1, \ldots, u_{n-1}\}$ supposed linearly independent, and called denominator-numerator estimation. This by analogy with linear systems terminology, where $q(s), p(s)$ is the numerator resp. the denominator of the transfer function $H(s)$.

3.3 Denominator estimation

The equation error is given by (21), hence the MSEE is

$$\|\tilde{e}_n\|^2 \quad = \quad \| \sum_{k=0}^{n} p_k v_k \|^2 \tag{22}$$

$$= \sum_{k=0}^{n}\sum_{l=0}^{n} p_k p_l^*(v_k, v_l) \tag{23}$$

$$= p^\dagger V p \tag{24}$$

where V is the $(n+1)\times(n+1)$ Hermitian positive semi-definite Grammian

$$V_{i,j} = (v_j, v_i) \quad \text{for} \quad i,j = 0,1,\ldots,n \tag{25}$$

and

$$p = (p_0, p_1, \ldots, p_n)^T \tag{26}$$

Taking the gradient of (24) with respect to p minimizes the MSEE straightforwardly. This yields $Vp = 0$, but the only solution of this equation is the trivial one, since in general $\det(V) \neq 0$, except when the order is exact and no noise is present. As pointed out in [12], there are three basic approaches to solve this kind of problem :

The Pencil-of-Function Method

The parameters p_i are given by

$$p_i = \sqrt{\Delta_{i,i}} \quad \text{for} \quad i = 0,1,\ldots,n \tag{27}$$

where $\Delta_{i,i}$ is the i-th diagonal cofactor of the matrix V.

Minimization with linear constraint

Minimize (24) with constraint $p_n = 1$. This yields equations similar to the discrete Wiener-Hopf equations.

Minimization with quadratic constraint

Minimize (24) with constraint $p^\dagger p = 1$. This is equivalent with finding the normalized eigenvector of V corresponding to the minimum eigenvalue.

Once the polynomial coefficients obtained, one can root $p(z)$ to yield the s_i. But what about the parameters c_i? The solution to this problem is to substitute the s_i in the MSE expression (6) and to minimize the MSE with respect to the c_i. This is a linear problem, since the MSE is a quadratic (but inhomogeneous) form in the c_i for any given set s_j. So

$$\|e_n\|^2 = \|v - \sum_{k=1}^{n} c_k y(s_k; 1)\|^2 \tag{28}$$

where the s_k are the already calculated roots of $p(z)$. It is readily seen that

$$c = (c_1, c_2, \ldots, c_n)^T \tag{29}$$

is the solution of the linear equation $Yc = h$, where

$$Y_{i,j} = (y(s_j;1), y(s_i;1)) \quad \text{for} \quad i,j = 1,2,\ldots,n \tag{30}$$

and

$$h_i = (v, y(s_i;1)) \quad \text{for} \quad i = 1,2,\ldots,n \tag{31}$$

3.4 Denominator-numerator estimation.

Define

$$\gamma = (\gamma_0, \gamma_1, \ldots, \gamma_{2n})^T \tag{32}$$

with

$$\gamma_{2k} = p_k \quad \text{for} \quad k = 0,1,\ldots,n \tag{33}$$

$$\gamma_{2k+1} = q_k \quad \text{for} \quad k = 0,1,\ldots,n-1 \tag{34}$$

This defines the polynomial

$$\gamma(z) = \sum_{k=0}^{2n} \gamma_k z^k = p(z^2) + zq(z^2) \tag{35}$$

Next define

$$\zeta_{2k} = v_k \quad \text{for} \quad k = 0,1,\ldots,n \tag{36}$$

$$\zeta_{2k+1} = -u_k \quad \text{for} \quad k = 0,1,\ldots,n-1 \tag{37}$$

Then (20) can be written as

$$\tilde{e}_n = \sum_{k=0}^{2n} \gamma_k \zeta_k \tag{38}$$

The MSEE is

$$\|\tilde{e}_n\|^2 = \|\sum_{k=0}^{2n} \gamma_k \zeta_k\|^2 \tag{39}$$

$$= \sum_{k=0}^{2n}\sum_{l=0}^{2n} \gamma_k \gamma_l^* (\zeta_k, \zeta_l) \tag{40}$$

$$= \gamma^\dagger G \gamma \tag{41}$$

where G is the $(2n+1) \times (2n+1)$ Hermitian positive semi-definite Grammian

$$G_{i,j} = (\zeta_j, \zeta_i) \quad \text{for} \quad i,j = 0,1,\ldots,2n \tag{42}$$

Again, in general $\det(G) \neq 0$ — the set $\{u_0, u_1, \ldots, u_{n-1}\}$ is supposed linearly independent — and a straightforward minimization of (41) yields the trivial solution. As in subsection 3.4 we therefore formulate possible solutions in terms of the three basic approaches.

The Pencil-of-Function Method

The parameters γ_i are given by

$$\gamma_i = \sqrt{\Delta_{i,i}} \tag{43}$$

where $\Delta_{i,i}$ is the i-th diagonal cofactor of the matrix G.

Minimization with linear constraint

Minimize the MSEE with constraint $p_n = \gamma_{2n} = 1$. This case permits an interesting imbedding interpretation. Putting $N = 2n$, the functional to be minimized is ($\gamma_N = 1$) :

$$J_N = \|\zeta_N + \sum_{k=0}^{N-1} \gamma_k \zeta_k\|^2 \tag{44}$$

Now if we suppose N to be free, the parameters γ_k which minimize (45) will evidently depend on N. Hence

$$J_N = \|\zeta_N + \sum_{k=0}^{N-1} \gamma_{N,k} \zeta_k\|^2 \tag{45}$$

It can be proved [14] that

$$\gamma_{N,k} = a_{N,k}/a_{N,N} \qquad J_N = 1/|a_{N,N}|^2 \tag{46}$$

where the $a_{i,j}$ are the coefficients of the Gram-Schmidt procedure transforming the set ζ_i into the orthonormal set ζ_i', i.e.,

$$\zeta_i' = \sum_{k=0}^{i} a_{i,k} \zeta_k \tag{47}$$

and

$$(\zeta_i', \zeta_j') = \delta_{i,j} \tag{48}$$

This is clearly an interesting computational procedure, since it is recursive, and with the $\gamma-$vector defined as above, updating $N \to N+2$ corresponds with $n \to n+1$. The Gram-Schmidt procedure can be shown (see Appendix B) to be equivalent with the Cholesky decomposition of the inverse of the Grammian G, what makes it possible to consider the use of such tools as fast Cholesky decomposition and lattice filters [16].

Minimization with quadratic constraint

The solution vector is the normalized eigenvector corresponding with the minimum eigenvalue of the Grammian G.

Note that the explicit functional form of $y(s;c)$ does not enter in the calculation of the Grammians G or V. All that is required is knowledge of the data-vector v, the operator K and the vector u. Moreover, the operator K plays a fundamental role in the construction of the Grammians G or V. If K is self-adjoint resp. unitary then the Grammians have a Hankel resp. Toeplitz structure, as proven in Appendix C.

4 APPLICATIONS

In what follows $H(\Omega)$ is the Hilbert space $L_2(a,b)$ with support $\Omega = [a,b]$ and scalar product

$$(x,y) = \int_a^b x(t)y^*(t)\,dt \tag{49}$$

The modeling approaches will be classified according to the support Ω, the operator K and the vector u.

4.1 Sinusoidal modeling.

$$\begin{aligned}
\Omega &= [-\infty,\infty] && (50)\\
y_t(s;c) &= ce^{j\omega t} \quad \omega \in R && (51)\\
K(t,t') &= \delta(t-t'+\Delta) && (52)\\
K^\dagger(t,t') &= \delta(t'-t+\Delta) && (53)\\
u &\equiv 0 && (54)\\
s &= e^{j\omega\Delta} && (55)
\end{aligned}$$

One sees that K is the forward shift operator, i.e.,

$$Kx(t) = x(t+\Delta) \tag{56}$$

and K is unitary, hence V will be Hermitian Toeplitz.

4.2 Exponential modeling.

$$\begin{aligned}
\Omega &= [0,\infty] && (57)\\
y_t(s;c) &= ce^{\alpha t} \quad \alpha \in C, \quad \Re(\alpha) \le 0 && (58)\\
K(t,t') &= \delta(t-t'+\Delta) && (59)\\
u &\equiv 0 && (60)\\
s &= e^{\alpha\Delta} && (61)
\end{aligned}$$

Again K is the forward shift operator. Note that $K^\dagger$ is not well defined with respect to the support Ω, so V is not Toeplitz in general.

4.3 Rational modeling.

$$\begin{aligned}
\Omega &= [0,T] && (62)\\
y_t(s;c) &= d\int_0^t e^{\alpha(t-t')}u_0(t')\,dt' && (63)\\
Kx &= \int_0^t x(t')\,dt' && (64)\\
K^\dagger x &= \int_t^T x(t')\,dt' && (65)\\
u &= \int_0^t u_o(t')\,dt' && (66)\\
s &= 1/\alpha && (67)\\
c &= -d/\alpha && (68)
\end{aligned}$$

The derivation of (62)-(68) is straightforward, based on the fact that output v and input u_0 of a rational system with transfer function

$$H(s) = \sum_{k=1}^{n} d_k/(s-\alpha_k) \tag{69}$$

are related by

$$v(t) = \left(\sum_{k=1}^{n} d_k e^{\alpha_k t}\right) * u_0(t) \tag{70}$$

where $*$ represents convolution. Note that u is the integrated input and that s in (67) is the inverse of the exponent (complex time constant).

The operator K has no evident nice properties, such as self-adjoint or unitary, which could reduce the computational burden associated with the Grammian G. However, introducing the reverse time integral [17], we can reformulate the problem in a more appropriate manner :

$$\begin{aligned} \Omega &= [0,\infty] && (71)\\ Kx &= \frac{1}{2}\left(\int_0^t x(t')\,dt' + \int_\infty^t x(t')\,dt'\right) && (72)\\ K^\dagger &= -K && (73)\\ u &= Ku_0 && (74)\\ s &= 1/\alpha && (75)\\ c &= -d/\alpha && (76) \end{aligned}$$

The fact that K is skew self-adjoint will reduce the number of scalar products to be calculated for the Grammian G from $O(n^2)$ to $O(n)$. This is valid only when

$$y_\infty(s;c) = 0 \tag{77}$$

for all s under consideration. This will be true when all poles are in the left half-plane and the input is e.g. time-limited.

4.4 Tauberian modeling.

As has been said before, Tauberian modeling is modeling by a weighted sum of shifted versions of the same known basis function x(t), i.e.,

$$v(t) = \sum_{k=1}^{n} c_k x(t - s_k) \quad s_k \in R \tag{78}$$

This means that, with the notation here

$$y_t(s;c) = cx(t-s) \tag{79}$$

In [11] the parameter extraction is performed using a Fourier-Prony approach, since transforming (78) yields

$$V(j\omega) = \sum_{k=1}^{n} c_k e^{-j\omega s_k} X(j\omega) \tag{80}$$

or

$$V(j\omega)/X(j\omega) = \sum_{k=1}^{n} c_k e^{-j\omega s_k} \tag{81}$$

which is a sum of complex exponentials, making application of Prony's method possible. In fact, referring to subsection 4.1, we can consider this problem as a sinusoidal modeling of the function $V(j\omega)/X(j\omega)$ in the frequency domain. Note that division by $X(j\omega)$ amounts to a hidden deconvolution in the time domain. In terms of the approach used here, we can alternatively formulate the problem as (see Appendix D)

$$\Omega = [-\infty, \infty] \tag{82}$$

$$K(t,t') = t\delta(t-t') + h(t-t') \tag{83}$$

$$u \equiv 0 \tag{84}$$

where the function $h(t)$ satisfies

$$h(t) * x(t) + tx(t) = 0 \tag{85}$$

with $*$ representing the double-sided convolution. Fourier-transforming (85) yields the filter

$$H(j\omega) = \frac{d}{dj\omega} \log X(j\omega) \tag{86}$$

The s_k can, as before, be extracted by processing appropriately the Grammian $V_{i,j} = (v_j, v_i)$, where $v_i = K^i v$, and rooting the polynomial $p(z)$. The problem which remains is whether the filter (86) is in some sense realizable.

4.5 Modeling by Bessel functions.

As a last illustrative example, we give modeling in terms of Bessel functions of order zero, i.e.,

$$v(t) = \sum_{k=1}^{n} d_k J_0(\alpha_k t) \tag{87}$$

Here $y_t(s;c) = dJ_0(\alpha t)$ and

$$\Omega = [0, T] \tag{88}$$

$$Kx = \int_0^t \frac{dt'}{t'} \int_0^{t'} \tau x(\tau)\, d\tau \tag{89}$$

$$u(t) = 1 \tag{90}$$

$$s = -1/\alpha^2 \tag{91}$$

$$c = d/\alpha^2 \tag{92}$$

Again, knowledge of the triplet (v, K, u) permits construction of the Grammian G and subsequent parameter extraction.

APPENDIX A

The MSE sequence is ever-decreasing.

$$\|e_n\|^2 = \|v - \sum_{k=1}^{n} y(s_k; c_k)\|^2 \tag{A1}$$

Suppose the minimum (there is one, since lower bound = 0) is reached when

$$s_k = s_{n,k}, \quad c_k = c_{n,k} \quad \text{for} \quad k = 1, 2, \ldots, n \tag{A2}$$

Then, for all values of s_k, c_k

$$\|v - \sum_{k=1}^{n} y(s_{n,k}; c_{n,k})\|^2 \leq \|v - \sum_{k=1}^{n} y(s_k; c_k)\|^2 \tag{A3}$$

But this must hold for the specific choice

$$c_n = 0, \quad s_k = s_{n-1,k}, \quad c_k = c_{n-1,k} \quad \text{for} \quad k = 1, 2, \ldots, n-1 \tag{A4}$$

and hence

$$\|e_n\|^2 \leq \|e_{n-1}\|^2 \tag{A5}$$

APPENDIX B

Gram-Schmidt orthonormalization and Cholesky decomposition.

Suppose the set ζ_i' is the result of the Gram-Schmidt orthonormalization performed on the set ζ_i, i.e.,

$$\zeta_i' = \sum_{k=0}^{i} a_{i,k}\zeta_k \tag{A6}$$

and

$$(\zeta_i', \zeta_j') = \delta_{i,j} \tag{A7}$$

Then

$$(\zeta_i', \zeta_j') = (\sum_{k=0}^{i} a_{i,k}\zeta_k, \sum_{l=0}^{j} a_{j,l}\zeta_l) = \delta_{i,j} \tag{A8}$$

Let A represent the lower triangular matrix $(a_{i,j})$. Then (A8) is equivalent with the matrix equation

$$AGA^{\dagger} = I \tag{A9}$$

or

$$G^{-1} = A^{\dagger}A \tag{A10}$$

where G is the Grammian associated with the ordered set ζ_i. But (A10) is precisely the Cholesky decomposition of G^{-1}.

APPENDIX C

Special structures for K and the Grammians V, G.

1. K is unitary : $K^{\dagger} = K^{-1}$

The Grammian V is Toeplitz :

$$V_{i,j} = (K^j v, K^i v) \tag{A11}$$

$$= (v, K^{\dagger j} K^i v) = (v, K^{i-j} v) \tag{A12}$$

Since Grammians in general exhibit Hermitian symmetry, we conclude that V is Hermitian Toeplitz. The number N_s of scalar products to be calculated is $N_s = n + 1$ and the Cholesky decomposition can be implemented using the Levinson-Durbin [4] algorithm. The rows and columns of the Grammian G can be rearranged to exhibit a block Toeplitz structure with $N_s = 4n + 1$.

2. K is self-adjoint (or skew self-adjoint) : $K^{\dagger} = K$ (or $K^{\dagger} = -K$)

The Grammian V has a Hankel structure :

$$V_{i,j} = (v, K^{i+j} v) \quad \text{if} \quad K^{\dagger} = K \tag{A13}$$

$$V_{i,j} = (-1)^j (v, K^{i+j} v) \quad \text{if} \quad K^{\dagger} = -K \tag{A14}$$

The factor $(-1)^j$ in (A14) has no effect on the number of scalar products to be calculated, which is $N_s = 2n + 1$. The rows and columns of the Grammian G can be rearranged to exhibit a block Hankel structure with $N_s = 6n$. We conclude that for K unitary, self-adjoint or skew, the number of scalar products to be calculated is $O(n)$ instead of $O(n^2)$.

APPENDIX D

Tauberian modeling.

We must show that

$$\int K(t,t')x(t'-s)\,dt' = sx(t-s) \tag{A15}$$

where $K(t,t')$ is defined in (83), (85). We have

$$\begin{aligned}\int K(t,t')x(t'-s)\,dt' &= tx(t-s) + \int h(t-t')x(t'-s)\,dt' &\text{(A16)}\\ &= tx(t-s) + \int h(t-t''-s)x(t'')\,dt'' &\text{(A17)}\\ &= tx(t-s) + (h*x)_{t-s} &\text{(A18)}\end{aligned}$$

But from (85) we obtain

$$(h*x)_t = -tx(t) \tag{A19}$$

and hence

$$\int K(t,t')x(t'-s)\,dt' = tx(t-s) - (t-s)x(t-s) = sx(t-s) \tag{A20}$$

References

[1] G. R. B. Prony, "Essai experimental et analytique, etc.," *Paris, J. de l'Ecole Polytechnique,* vol. 1, cahier 2, pp. 24–76, 1795.

[2] C. E. Baum, "Emerging technology for transient and broadband analysis and synthesis of antennas and scatterers," *Interaction Notes,* 300, Nov. 1976; also in *Proc. IEEE ,* vol. 64, pp. 1598–1676, Nov. 1976.

[3] M. L. Van Blaricum and R. Mittra, "A technique for extracting the poles and residues of a system directly from its transient response," *Interaction Notes,* 245, Feb. 1975; also in *IEEE Trans. Antennas Propagat.,* vol. AP-23, Nov. 1975.

[4] S. M. Kay and S. L. Marple, Jr., "Spectrum analysis — a modern perspective," *Proc. IEEE,* vol. 69, pp. 1380–1419, Nov. 1981.

[5] T. K. Sarkar, D. D. Weiner, J. Nebat, and V. K. Jain, "A discussion of various approaches to the identification/approximation problem," *IEEE Trans. Antennas Propagat.,* vol. AP-30, Jan. 1982.

[6] T. K. Sarkar, J. Nebat, D. D. Weiner, and V. K. Jain, "Suboptimal approximation/identification of transient waveforms from electromagnetic systems by Pencil-of-Function method," *IEEE Trans. Antennas Propagat.,* vol. AP-28, Nov. 1980.

[7] F. Brophy and A. C. Salazar, " Considerations of the Padé approximant technique in the synthesis of recursive digital filters," *IEEE Trans. Audio and Electroacoustics,* vol. AU-21, pp. 500–505, Dec. 1973.

[8] R. N. McDonough and W. H. Huggins, " Best least-squares representation of signals by exponentials," *IEEE Trans. Automat. Contr.,* vol. AC-13, pp. 408–412, Aug. 1968.

[9] C. B. Sharpe and C. J. Roussi, "Equivalent circuit representation of radiation systems," *AFWL Interaction Note,* 361, April 1979.

[10] D. G. Dudley, "Parametric modeling of transient electromagnetic systems," *Radio Science* vol. 14, no. 3, pp. 387–396, May-June 1979.

[11] R. J. P. de Figueredo and C. L. Hu, "Waveform feature extraction based on Tauberian approximation," *IEEE Trans. Pattern Anal. Machine Intell.,* vol. PAMI-4, no. 2, pp. 105–116, March 1982.

[12] T. K. Sarkar, S. A. Dianat, and D. D. Weiner, "A discussion of various approaches to the linear system identification problem," *IEEE Trans. Acoust., Speech, Signal Processing,* vol. ASSP-32, no. 3, pp. 654–656, June 1984.

[13] R. E. Kalman, "Design of a self-optimizing control system," *Trans. ASME,* vol. 80, pp. 468–474, Feb 1958.

[14] P. J. Davis, *Interpolation and Approximation,* New York : Dover, 1975.

[15] J. Makhoul, "Linear prediction : A tutorial review," *Proc. IEEE,* vol. 63, pp. 561–580, April 1980.

[16] M. Morf, "Fast algorithms for multivariable systems," *Ph. D. dissertation, Dept. Elec. Eng.,* Stanford Univ., Stanford, CA, 1974.

[17] V. K. Jain, "Filter analysis by use of pencil functions : Part I," *IEEE Trans. Circuits Syst.,* vol. CAS-21, Sept. 1974.