Estimation of parameters
of linear systems
using periodic test signals

Estimation of parameters of linear systems using periodic test signals

PROEFSCHRIFT *ter verkrijging van de graad van doctor in de technische wetenschappen aan de Technische Hogeschool Delft, op gezag van de rector magnificus ir. H. B. Boerema, hoogleraar in de afdeling der elektrotechniek, voor een commissie aangewezen door het college van dekanen te verdedigen op woensdag 26 juni 1974 te 14.00 uur door*

ADRIAAN VAN DEN BOS

natuurkundig ingenieur, geboren te Rotterdam

1974 / Delftse Universitaire Pers

Dit proefschrift is goedgekeurd
door de promotor
PROF. IR. B. P. TH. VELTMAN

Aan mijn ouders
Aan mijn schoonouders

SUMMARY

In this thesis a procedure is developed for estimating parameters of linear systems from noise corrupted responses to periodic test signals. The class of linear systems is restricted to linear systems describable by an ordinary linear differential equation. The parameters to be estimated are the coefficients of the differential equation and a time delay in series with the system. The proposed procedure is a weighted least squares procedure operating on the estimates of the Fourier coefficients of the test signal and those of the response. The estimator of the coefficients of the differential equation is a simple closed form expression in the estimator of the Fourier coefficients. The estimates are obtained in a single computational step. No iterations are required. The time delay is estimated by repeating the procedure for a number of values of time delay and selecting the best fitting solution. The proposed procedure is consistent if the following conditions are both satisfied: 1) the number of unknown parameters may not exceed twice the number of harmonics taken into consideration and 2) the covariance function of the noise is absolutely integrable. The procedure is applicable to systems under closed loop control.

An expression is derived for the covariance matrix of the proposed estimator. This expression shows how the weights of the least squares procedure must be chosen in order to minimize the variance. These particular weights will be referred to as optimal weights. It is shown that the minimum variance coincides with the minimum variance bound (Cramér-Rao lower bound) if the noise is normally distributed. The optimal weights are functions of the properties of system and noise and are therefore not known

a priori. In order to obtain an estimator having approximately the minimum variance a second computational step can be added to the proposed procedure. In the first step the parameters are estimated using uniform weights. Using the results of the first step the optimal weights are estimated. These estimated weights are used as weights in the second step. Numerical results obtained from computer generated data show the close agreement of the variance of the two-step procedure with the minimum variance bound in the cases considered.

It is shown that the elements of the minimum variance bound can be expressed as functions of the power spectrum of the test signal and of the dynamics of system and noise. So for given system and noise these elements can be manipulated by selecting the power spectrum of the test signal. Numerically a number of test signal power spectra have been computed which minimize the trace of the minimum variance bound. The aim is to obtain a reference to which the minimum variance bound computed for the usual test signals can be compared. A further aim is to investigate how a priori knowledge about system and noise may be utilized for selection of appropriate test signals.

Finally, a numerical procedure is developed for approximate design of periodic two-level test signals having specified spectra. Numerical examples of signals computed using this procedure are described.

SAMENVATTING

Dit proefschrift beschrijft het ontwerp van een schattingsprocedure met behulp waarvan parameters van lineaire systemen kunnen worden bepaald uit door ruis verstoorde responsies op periodieke testsignalen. De beschouwde klasse lineaire systemen omvat uitsluitend die lineaire systemen die kunnen worden beschreven met behulp van een gewone lineaire differentiaalvergelijking. De te schatten parameters zijn de coefficienten van de differentiaalvergelijking en een voortplantingstijd in serie met het systeem. De schattingsprocedure is een gewogen kleinste-kwadratenmethode die wordt toegepast op de geschatte fouriercoefficienten van het testsignaal en die van de responsie. De schatter van de coefficienten van de differentiaalvergelijking is een eenvoudige expliciete uitdrukking in de schatter van de fouriercoefficienten. De schatting wordt verkregen in één enkele stap. Er behoeft niet te worden geitereerd. De voortplantingstijd wordt geschat door de procedure te herhalen voor een aantal waarden van de voortplantingstijd en vervolgens de best passende oplossing te bepalen. De schattingsprocedure is asymptotisch raak als aan de volgende voorwaarden is voldaan: 1) het aantal te schatten parameters mag niet groter zijn dan tweemaal het aantal harmonischen dat bij de schatting in aanmerking wordt genomen en 2) de covariantiefunctie van de ruis is absoluut integreerbaar. De schattingsprocedure mag ook worden toegepast op systemen opgenomen in een regellus.

Een uitdrukking voor de covariantiematrix van de schatter wordt afgeleid. Uit deze uitdrukking blijkt hoe de weegfactoren van de kleinste-kwadratenmethode gekozen moeten worden om de variantie te minimaliseren. Deze weegfactoren worden in het volgende optimale weegfactoren genoemd. Aan-

getoond wordt dat de minimale variantie samenvalt met de cramér-raobenedengrens als de ruis normaal verdeeld is. De optimale weegfactoren zijn een functie van de eigenschappen van het systeem en de ruis en zijn dan ook niet a priori bekend. Om een schatter te verkrijgen waarvan de variantie de minimale variantie benadert kan de schattingsprocedure worden uitgebreid met een tweede stap. In de eerste stap worden de systeemparameters geschat waarbij de gewichtsfactoren van de kleinste-kwadraten procedure onderling gelijk gekozen zijn. Vervolgens worden de optimale gewichtsfactoren geschat met behulp van de uitkomsten van de eerste stap. Deze geschatte gewichtsfactoren worden als gewichtsfactoren gebruikt in de tweede stap. Deze twee-stappen procedure is toegepast op gegevens die werden gegenereerd met behulp van een digitale rekenautomaat. De variantie van de twee-stappen procedure blijkt in de beschouwde gevallen met de cramér-raobenedengrens overeen te komen.

De elementen van de cramér-raobenedengrens blijken te kunnen worden geschreven als functies van het vermogensdichtheidsspectrum van het testsignaal en de dynamische eigenschappen van het systeem en de ruis. Bij een gegeven systeem en ruis kan dan de grootte van deze elementen worden beinvloed door de keuze van het vermogensdichtheidsspectrum van het testsignaal. Numeriek zijn een aantal spectra berekend die het spoor van de cramér-raobenedengrens minimaliseren. Het doel van deze berekeningen is een referentie te verkrijgen waarmee de cramér-raobenedengrens berekend voor de gebruikelijke testsignalen kan worden vergeleken. Verder hebben de berekeningen tot doel na te gaan hoe a priori kennis van het systeem en de ruis kan worden benut bij de keuze van een testsignaal.

Tenslotte wordt een numerieke procedure beschreven met behulp waarvan twee-standensignalen kunnen worden ontworpen die bij benadering een voorgeschreven vermogensdichtheids-spectrum bezitten. Numerieke voorbeelden worden gegeven van signalen die zijn berekend met behulp van deze procedure.

CONTENTS

LIST OF SYMBOLS AND ABBREVIATIONS

(Number in parentheses refers to page where the quantity is defined or first introduced)

a_n	coefficient of nth order derivative in left-hand member of system differential equation (21)
b_m	coefficient of mth order derivative in right-hand member of system differential equation (21)
c	vector of a_n and b_m coefficients (22)
$\hat{c}^{(J)}$	least squares estimator of c (25)
d	vector (22)
$\hat{d}^{(J)}$	least squares estimator of d (22)
$\hat{d}_w^{(J)}$	vector valued least squares estimator (111)
$e(t)$	(in Section 1.4) zero-mean, normal process (44) (in Section 3.5) zero-mean white, normal process(104)
$f(t)$	(in Section 1.4) input of controller (50) (in Section 3.1) periodic signal (79)
$g(t)$	additive noise in observations of u(t) (24)
$g_i(t)$	component of g(t) (32)
$h(t)$	additive noise in observations of y(t) (24)
$h_i(t)$	component of h(t) (32)
$h_S(t)$	system impulse response (102)
i	integer
i_o	uniformly distributed, integer, random number (70)

i_1	uniformly distributed, integer, random number (70)
j	$\sqrt{-1}$
k	integer
k_h	integer (42)
k_i	harmonic number (22)
k_w	integer denoting half-width of spectral window (40)
ℓ	integer
m	integer
n	integer
$n(t)$	stationary stochastic process (80)
p	integer
p lim	probability limit (87)
q	integer
s	$j\omega$
s_k	$j2\pi k/T$ (21)
$s_{k/J}$	$j2\pi k/JT$ (110)
t	time
t_1	time
$u(t)$	periodic input (21)
$u(i)$	value of discrete interval, binary, periodic signal on ith interval (69)
$v(t)$	observations of $u(t)$ (24)

$v_{q_1q_2,\check{\eta}\check{\eta}}$	q_1q_2 element of $V_{\check{\eta}\check{\eta}}$ (131)
$w(t)$	sum of periodic signal and stochastic process (79)
x_{pq}	pq element of X (55)
$x_{pq,mn}$	mn element of X_{pq} (59)
y	vector (58)
y_ℓ	element of y (58)
$y(t)$	steady state response to $u(t)$ (21)
$z(t)$	observations of $y(t)$ (24)
$\bar{z}(t)$	$z(t)$ after subtracting moving average (30)

$A(j\omega)$	denominator polynomial of system transfer function (25)
$B(j\omega)$	numerator polynomial of system transfer function (25)
$B_\tau(j\omega)$	$B(j\omega)\exp(-j\omega\tau)$ (95)
D	weighted sum of diagonal elements of $\Sigma_{\tilde{\theta}\tilde{\theta}}$ (57)
D_k	matrix (96)
$D(j\omega)$	polynomial in $j\omega$ (85)
E	expectation operator
$E(j\omega)$	polynomial in $j\omega$ (85)
$F(j\omega)$	transfer function (107)

$F_i(j\omega)$	transfer function (100)
$\mathscr{F}$	Fourier integral operator (103)
G	matrix (94)
G_k	matrix (96)
$G_{i/J}$	matrix (119)
G_W	matrix (118)
$G_{W'}$	matrix (123)
$H_S(j\omega)$	system transfer function (55)
$H_D(j\omega)$	noise transfer function (104)
$H_i(j\omega)$	transfer function (56)
I	identity matrix (37)
I_s	number of steps in a period of discrete interval, binary, periodic signal (69)
I_η	dimension of $\check{\eta}^{(J)}$ (129)
I_θ	dimension of $\check{\theta}^{(J)}$ (129)
J	integer denoting number of observed periods (24)
L	number of harmonics taken into consideration (22) (in Section 3.5) likelihood function (104)
M	order of right-hand member of system differential equation (21)
N	order of left-hand member of system differential equation (21)

O	matrix or part of matrix consisting of zeros only
O_1	zero matrix (118)
O_2	zero matrix (118)
P	matrix (22)
$\hat{P}^{(J)}$	estimator of P (25)
P_k	matrix (22)
$\hat{P}_w^{(J)}$	matrix (111)
$\hat{P}_{k/J,w}^{(J)}$	matrix (110)
$R_{gg}(t)$	autocovariance function of g(t) (24)
$R_{gh}(t)$	cross covariance function of g(t) and h(t) (33)
$R_{hh}(t)$	autocovariance function of h(t) (24)
$R_{nn}(t)$	autocovariance function of n(t) (80)
$S(j\omega_{k_\ell})$	desired value of $\|\gamma_{k_\ell}\|^2$ (69)
$S_{gg}(j\omega)$	power density spectrum of g(t) (33)
$S_{hh}(j\omega)$	power density spectrum of h(t) (33)
$S_{uu}(j\omega_k)$	$\|\gamma_{ku}\|^2$ (55)
$\hat{S}_{uu}(j\omega_k)$	optimal value of $S_{uu}(j\omega_k)$
S_o	constant (61)
T	period of periodic (test) signal (21)
T'	period of periodic process (82)

T_{obs}	observation time (63)
$V(k)$	finite Fourier transform of $v(t)$ (37)
$V_\tau(k)$	$V(k)\exp(-s_{k/J}\tau)$ (110)
$V_{\hat{\eta}\hat{\eta}}$	asymptotic covariance matrix of $\sqrt{JT}\,\hat{\eta}^{(J)}$ (90)
$V_{\check{\eta}\check{\eta}}$	asymptotic covariance matrix of $\sqrt{J}\,\check{\eta}^{(J)}$ (129)
$V_{\hat{\eta}_c\hat{\eta}_c}$	asymptotic covariance matrix of $\sqrt{JT}\,\hat{\eta}_c^{(J)}$ (114)
$V_{\hat{\eta}_w\hat{\eta}_w}$	asymptotic covariance matrix of $\sqrt{JT}\,\hat{\eta}_w^{(J)}$ (122)
$V_{\hat{\theta}\hat{\theta}}$	asymptotic covariance matrix of $\sqrt{JT}\,\hat{\theta}^{(J)}$ (32)
$V_{\check{\theta}\check{\theta}}$	asymptotic covariance matrix of $\sqrt{J}\,\check{\theta}^{(J)}$ (132)
W_k	matrix (92)
$W_{i/J}$	matrix (117)
X	matrix (55)
X_{pq}	matrix obtained by eliminating pth row and qth column from X (58)
$X_{pq,\,mn}$	matrix obtained by eliminating mth row and nth column from X_{pq} (59)
$Z(k)$	finite Fourier transform of $z(t)$ (37)
α_{kf}, α_{ku}, α_{ky}	Fourier cosine coefficients of kth harmonic of periodic signals $f(t)$, $u(t)$ and $y(t)$ respectively (21)

$\hat{\alpha}^{(J)}_{kf}$, $\hat{\alpha}^{(J)}_{ku}$, $\hat{\alpha}^{(J)}_{ky}$	least squares estimators of α_{kf}, α_{ku} and α_{ky} respectively (24)
β_{kf}, β_{ku}, β_{ky}	Fourier sine coefficients of kth harmonic of periodic signals f(t), u(t) and y(t) respectively (21)
$\hat{\beta}^{(J)}_{kf}$, $\hat{\beta}^{(J)}_{ku}$, $\hat{\beta}^{(J)}_{ky}$	least squares estimators of β_{kf}, β_{ku} and β_{ky} respectively (24)
γ_{k_ℓ}	complex Fourier coefficient of kth harmonic of descrete interval, binary, periodic signal (69)
γ_{kf}, γ_{ku}, γ_{ky}	complex Fourier coefficients of kth harmonic of periodic signals f(t), u(t) and y(t) respectively (21)
$\hat{\gamma}^{(J)}_{kf}$, $\hat{\gamma}^{(J)}_{ku}$, $\hat{\gamma}^{(J)}_{ky}$	least squares estimators of γ_{kf}, γ_{ku} and γ_{ky} respectively (24)
$\delta(t)$	Dirac delta function
$\hat{\varepsilon}^{(J)}$	vector of residuals (25)
$\hat{\varepsilon}^{(J)}_{w}$	vector of residuals (111)
$\hat{\varepsilon}^{(J)}_{w\hat{\theta}}$	vector of measured residuals (112)
$\hat{\varepsilon}^{(J)}_{w'}$	vector of residuals (124)
$\hat{\varepsilon}^{(J)}_{w'\hat{\theta}}$	vector of measured residuals (124)
η	vector of Fourier coefficients of u(t) and y(t) (90)
η_c	(asymptotic) expectation of $\hat{\eta}^{(J)}_{c}$ (114)

η_w	(asymptotic) expectation of $\hat{\eta}_w^{(J)}$ (118)
$\eta_{w'}$	(asymptotic) expectation of $\hat{\eta}_{w'}^{(J)}$ (124)
$\check{\eta}$	asymptotic expectation of $\check{\eta}^{(J)}$ (129)
$\check{\eta}_q$	qth element of $\check{\eta}$ (129)
$\hat{\eta}^{(J)}$	least squares estimator of η (90)
$\check{\eta}^{(J)}$	arbitrary estimator (129)
$\hat{\eta}_c^{(J)}$	vector of finite Fourier transforms of v(t) and z(t) (113)
$\hat{\eta}_w^{(J)}$	vector of finite Fourier transforms of v(t) and z(t) (112)
$\hat{\eta}_{w'}^{(J)}$	vector of finite Fourier transforms of v(t) and z(t) (124)
θ	vector of system parameters (23) (in Section 3.5) vector of system and noise parameters (104)
θ_S	(in Section 3.5) vector of system parameters (102)
θ_D	(in Section 3.5) vector of noise parameters (104)
$\check{\theta}$	asymptotic expectation of $\check{\theta}^{(J)}$ (129)
$\check{\theta}_p$	pth element of $\check{\theta}$ (129)
$\hat{\theta}^{(J)}$	least squares estimator of θ (32)
$\hat{\theta}_I^{(J)}$	least squares estimator of θ using weighting matrix I (38)
$\hat{\hat{\theta}}^{(J)}$	two-step least squares estimator of θ (40)
$\tilde{\theta}^{(J)}$	minimum variance least squares estimator of θ (35)

$\breve{\theta}^{(J)}$	arbitrary estimator (129)
$\breve{\theta}(\breve{\eta}^{(J)})$	$\breve{\theta}^{(J)}$ (129)
$\breve{\theta}_p^{(J)}$	pth element of $\breve{\theta}^{(J)}$ (130)
$\breve{\theta}_p(\breve{\eta}^{(J)})$	pth element of $\breve{\theta}(\breve{\eta}^{(J)})$ (129)
$\lambda(\omega_k)$	diagonal element of Λ (34)
$\hat{\lambda}_w(\omega_k)$	estimator of $\lambda(\omega_k)$ (40)
$\hat{\lambda}_{w'}(\omega_k)$	estimator of $\lambda(\omega_k)$ (42)
λ_ℓ	positive weighting factor (57)
μ	expectation of n(t) (80)
μ_n	(in Section 2.1) positive weighting factor (57) (in Section 3.2) scalar (85)
ν_m	scalar (85)
π_{pq}	pq element of Π (104)
ρ_i	coefficient of trend polynomial (29)
σ	standard deviation (65)
σ_{ij}	ij element of $\Sigma_{\tilde{\theta}\tilde{\theta}}$ (57)
τ	time delay (21)
$\hat{\tau}^{(J)}$	least squares estimator of τ (26)
φ	phase angle (74)
ψ_{pq}	pq element of Ψ (107)
ω	frequency in rad s^{-1}

ω_k	$2\pi k/T$
ω_1	fundamental frequency in rad s^{-1}
Λ	diagonal matrix (34)
$\hat{\Lambda}$	estimator of Λ (40)
Π	information matrix (104)
$\Sigma_{\hat{\varepsilon}\hat{\varepsilon}}$	asymptotic covariance matrix of $\hat{\varepsilon}^{(J)}$ (35)
$\Sigma_{\hat{\varepsilon}_w\hat{\varepsilon}_w}$	asymptotic covariance matrix of $\hat{\varepsilon}_w^{(J)}$ (122)
$\Sigma_{\hat{\varepsilon}_{w'}\hat{\varepsilon}_{w'}}$	asymptotic covariance matrix of $\hat{\varepsilon}_{w'}^{(J)}$ (124)
$\Sigma_{\hat{\varepsilon}_{w\hat{\theta}}\hat{\varepsilon}_{w\hat{\theta}}}$	asymptotic covariance matrix of $\hat{\varepsilon}_{w\hat{\theta}}^{(J)}$ (117)
$\Sigma_{\hat{\varepsilon}_{w'\hat{\theta}}\hat{\varepsilon}_{w'\hat{\theta}}}$	asymptotic covariance matrix of $\hat{\varepsilon}_{w'\hat{\theta}}^{(J)}$ (125)
$\Sigma_{\hat{\eta}\hat{\eta}}$	asymptotic covariance matrix of $\hat{\eta}^{(J)}$ (90)
$\Sigma_{\check{\eta}\check{\eta}}$	asymptotic covariance matrix of $\check{\eta}^{(J)}$ (129)
$\Sigma_{\hat{\eta}_c\hat{\eta}_c}$	asymptotic covariance matrix of $\hat{\eta}_c^{(J)}$ (114)
$\Sigma_{\hat{\eta}_w\hat{\eta}_w}$	asymptotic covariance matrix of $\hat{\eta}_w^{(J)}$ (122)
$\Sigma_{\hat{\theta}\hat{\theta}}$	asymptotic covariance matrix of $\hat{\theta}^{(J)}$ (32)
$\Sigma_{\tilde{\theta}\tilde{\theta}}$	asymptotic covariance matrix of $\tilde{\theta}^{(J)}$ (35)
Φ	matrix (34)

Φ_i matrix (34)

$\Phi_{i/J}$ matrix (120)

Φ_W matrix (120)

Ψ information matrix (106)

Ω^{-1} positive definite weighting matrix (25)

DFT discrete Fourier transform (28)

FFT fast Fourier transform (28)

GLS generalized least squares (10)

IV instrumental variable(s) (10)

MFBS multifrequency binary signal (43)

MFBS1 particular MFBS (43)

MFBS2 particular MFBS (43)

ML maximum likelihood (10)

MLBS maximum length binary sequence (43)

Estimation of parameters
of linear systems
using periodic test signals

INTRODUCTION

The subject of this thesis belongs to the field of identification and system parameter estimation. Identification and system parameter estimation techniques are employed in many areas of research. Important applications are:

- Estimation of parameters of dynamical systems for improved control;
- Investigation of dynamical properties of mechanical systems for design purposes;
- Estimation of physical properties of materials using dynamical methods;
- Study of dynamical relations in biological systems.

For a survey of identification and system parameter estimation techniques the reader is referred to Åström and Eykhoff (1971). In what follows only a few of these methods will briefly be discussed for comparison purposes.

The main topic discussed in this thesis is a method developed by the author for estimating the parameters of a particular class of linear dynamical systems from noise corrupted observations of periodic input-output pairs. A convenient method to characterize identification and system parameter estimation problems is to specify the class of models, the class of inputs and the criterion of equivalence of the system under test and the elements of the class of models. This classification was first introduced by Zadeh (1962) and is also discussed by Åström and Eykhoff (1971). Using this method the identification problem discussed in this thesis may be characterized as follows.

The class of models is restricted to models describable by an ordinary linear scalar differential equation with constant coefficients. Within this class models having an unknown, time invariant time delay in series are also allowed. So the class of models is a parametric class, the parameters being the coefficients of the differential equation and the time delay. For simplicity the coefficients of the differential equation will be referred to as the system coefficients.

The most important restriction on the class of inputs is the assumption that the test signal is periodic. More precisely, it is assumed that the input of the system has a periodic component of given fundamental period. Two distinct cases will be considered. In Case 1 the system input has a completely known periodic component. This component may be the test signal itself or it is a signal having a known linear, dynamic or static relation with the test signal. For example, the latter situation occurs if the test signal is transformed by a linear transducer having known linear dynamics. In Case 2 the system input has a periodic component which is only partly known. The periodic component has an unknown linear, static or dynamic relation with the test signal. This is the case of the test signal is applied to the system through a linear transducer having unknown linear dynamics. Case 2 also includes the important case that the system is under closed loop control. For example, if the test signal is introduced at the set point, the periodic signal entering the system is not the test signal itself. It is a signal having an unknown dynamic relation with the test signal. The most important difference between Case 1 and Case 2 is that, although in both cases the period is given, in the latter case the periodic component of the input must be measured. It will be assumed that these measurements are corrupted by additive noise. For example, if the system is under closed loop control the noise may represent normal operating signals. Both in Case 1 and Case 2 it will be assumed that the observations at the output are additive noise corrupted observations of the steady state response to the periodic component of the input. Finally, it is assumed that both in Case 1 and Case 2 the periodic component of the input contains a sufficiently large number of non-zero harmonics. This number will be specified later.

From the above specification of the input it is clear that in this thesis the use of spontaneous statistical fluctuations for identification purposes will not be considered. An advantage of the use of spontaneous fluctuations is that one needs not to disturb the normal operation of the system. A disadvantage of spontaneous fluctuations is that their power spectrum cannot be selected and may be inappropriate for identification purposes. Generally the properties of the power spectrum of the input have a substantial influence on the accuracy in estimating the dynamical properties of the system. Furthermore input output observations made on spontaneous fluctuations may have covariant components which are not causally related by the system. For example, these components may be related through a, possibly hidden, feedback path. In the case of non-causal covariation of input and output some identification methods produce wrong results, since these methods interpret all covariation as causal. A further disadvantage of spontaneous fluctuations is that they are often nonstationary. Nonstationarity precludes the use of most identification methods. Finally, many spontaneous processes have limited power. Consequently observations are subject to relatively large measuring errors. In conclusion, it is preferable to employ external, artificial test signals whenever this is allowed and possible.

The restriction of the class of inputs to periodic ones also excludes external transient test signals such as pulses. In practice transient test signals are very useful for pilot estimation or for estimation in situations where noise is virtually absent. For estimation in the presence of noise, however, transient test signals are less suitable, since they make a very uneconomic use of the allowable input amplitude range. In practice the input amplitude is always restricted. Too large an amplitude may drive the system into non-linearity; usually a linear model is a small signal model. Moreover, too large an amplitude may disturb the normal operation too much. Furthermore the allowable amplitude range of the input transducer is usually limited. It is true that the signal-to-noise ratio in the observations of the response can be inproved by average response techniques. But this is equivalent to employing a periodic test signal. In that case, however, alternative periodic test signals of the same maximum amplitude are available which yield results having higher statistical accuracy. Particular examples of such signals discussed in this thesis are maximum length binary sequences and multifrequency binary signals. Maximum length binary sequences have their power evenly distributed over many harmonics. As opposed to maximum length binary sequences multifrequency binary signals have the major part of their power concentrated in a limited number of relatively widely spaced harmonics. The design of multifrequency binary signals is discussed in Section 2.3 of this thesis. For a discussion of maximum length binary sequences the reader is referred to Hoffmann de Visme (1971).

Finally, the above characterisation of the identification problem requires the definition of a criterion of equivalence of the system under test and the elements of the class of models. This criterion is defined as follows. The periodic component of the input and the corresponding periodic steady state response component of the output satisfy the differential equation of the system. So the Fourier coefficients of these components satisfy the Fourier transform of the differential equation. For each harmonic this yields two linear algebraic equations in the system coefficients. These algebraic equations will be referred to as the system equations. In the case considered in this thesis the observations of the periodic input component and those of the corresponding response component are corrupted by noise. The Fourier coefficients of these components can therefore only be estimated. Now suppose that estimates of a number of Fourier coefficients of the periodic components of input and output are available. Define the residual of a system equation as the difference between the left hand and the right hand member of the equation after replacing the Fourier coefficients by their estimates. Then the criterion of equivalence chosen in this research is a positive quadratic form in the residuals of all harmonics taken into consideration. The selection of the weighting matrix corresponding to this quadratic form will be discussed later. This definition of the criterion of equivalence of the system under test and the elements of the class of models completes the characterization of the identification problem studied in this thesis.

From the above considerations it is clear that for known order of the differential equation the problem is to minimize the criterion with respect to the unknown parameters. First consider the case that the time delay is known. Since the residuals are linear in the unknown system coefficients, the criterion is quadratic in these coefficients. The minimization of the criterion therefore only requires the solution of a set of linear algebraic equations. This is a closed form, one step solution. No iterations are required. The minimum is always achieved and this minimum is unique, provided that the normal equations are linearly independent. In Section 3.2 it is shown that this is always the case if the number of unknown system coefficients does not exceed twice the number of non-zero harmonics taken into consideration. In the case of unknown time delay the procedure is repeated for a number of selected values of time delay and the optimal value of time delay and the corresponding solution for the system coefficients are selected. Since the test signal is periodic, time delay can only be estimated modulo the period of the fundamental of the test signal. It is therefore assumed that the time delay consists of the sum of a known integral multiple of periods of the fundamental and an unknown fraction of this period. For unknown system order the above procedure is repeated for different orders in order to find the optimal solution.

In Section 3.2 it is shown that the estimation procedure described above is consistent if the following conditions are both satisfied: 1) the estimator of the Fourier coefficients is consistent and 2) the number of unknown system parameters does not exceed twice the number of harmonics taken into consideration.

The estimator of the Fourier coefficients chosen in this research is the least squares estimator. The main motives for this choice are that this estimator is computationally convenient and that it requires little a priori knowledge about the noise. Relevant statistical properties of this estimator are discussed in Section 3.1. Sufficient conditions for its mean square convergence are: 1) the noise is a stationary stochastic process and 2) the covariance function of this process is absolutely integrable. The former condition requires that non-stationarities are removed. An example of a non-stationarity frequently encountered in practice is a trend. Schemes for elimination of trends are discussed in Section 1.1. The latter condition is met by all processes having rational power spectra. No assumptions are made with respect to the amplitude distribution of the noise.

The second condition for the consistency of the estimator of the system parameters answers the question how many harmonics are required. The problem which harmonics and what fundamental frequency must be chosen will be discussed below.

A substantial part of this thesis is devoted to an approximate computation of the covariance matrix of the proposed estimator. See Section 3.4. This computation is motivated by the following questions: 1) how should the weights in the criterion be chosen for minimum variance?, 2) how large is the variance of the proposed estimator compared to the minimum variance bound? and 3) what is the influence of the spectrum of the test signal upon the elements of the covariance matrix of the proposed estimator? These questions will now first be discussed.

Since the estimator of the system parameters is non-linear in the estimators of the Fourier coefficients, closed form expressions for its small sample covariance matrix are hard to obtain. The asymptotic expression for this covariance matrix is, however, relatively easy to compute, provided that for orders higher than two the central moments of the estimators of the Fourier coefficients are more than inversely proportional to the observation time. The expression for the asymptotic covariance matrix shows that this matrix is smallest if the weighting matrix in the criterion is the inverse of the asymptotic covariance matrix of the residuals. It is shown in Section 3.4 that the latter covariance matrix is diagonal. The estimator of the system parameters corresponding to the optimal weighting matrix will be referred to as the minimum variance least squares estimator. In practice the variances of the residuals are not known and must therefore be estimated from the observations. A two-step procedure, described in Section 1.3, computes in the first step the system parameters using the identity matrix as weighting matrix and next estimates the variances of the residuals. The reciprocal values of the estimated variances are used as weights in the second step. This estimator will be referred to as the two-step least squares estimator. Numerical examples described in Section 1.4 show that at least in the cases considered the two-step procedure achieves the variance of the minimum variance least squares estimator. For what follows it is important to note that in these examples the open loop case is considered and that the noise has a normal distribution.

The minimum variance bound for the estimation of the system parameters is computed in Section 3.5. The case considered is the open loop case. The noise obeys the normal distribution. The expression for the minimum variance bound is identical to the expression for the asymptotic covariance matrix of the minimum variance least squares estimator. The important conclusion to be drawn from this identity is that asymptotically the minimum variance least squares estimator achieves the minimum variance bound if the noise is normal. So in the above-mentioned numerical examples of the two-step procedure, the minimum variance bound is actually achieved.

The expression for the minimum variance bound shows that the only property of the test signal affecting the elements of this bound is the power density spectrum of the test signal. The elements of the inverse of the minimum variance bound are relatively simple linear expressions in the power of the harmonics. The coefficients of these expressions are functions of the dynamic properties of the system and those of the noise. Observing that in practice input power or output power is always restricted, the question then arises which spectrum gives the most accurate results under this constraint. As a measure of accuracy the weighted sum of the diagonal elements of the minimum variance bound is chosen. The spectrum and the corresponding test signals which minimize this measure are defined as optimal. Section 2.1 describes the numerical procedure for mimimimization of the measure. In Section 2.2 some numerically computed optimal spectra are presented. It is striking that these spectra consist of a very limited number of harmonics only. For comparison purposes Section 2.2 also computes the minimum variance bound for a maximum length binary sequence and for an arbitrarily chosen multifrequency binary signal. These computations show that in a limited range of the fundamental frequency the variances with the maximum length binary sequence and the multifrequency binary signal are comparable to the variances with the optimal signals. This indicates that in the cases considered an appropriate bandwidth is more important than the specific shape of the spectrum within this bandwidth. The bandwidth of a particular maximum length binary sequence or a multifrequency binary signal is determined by its fundamental frequency. In order to find an appropriate fundamental frequency for an experiment it seems worthwhile to carry out a priori computations of the minimum variance bound for a number of different fundamental frequencies. These computations may be based on a priori knowledge about system and noise obtained from mathematico-physical analysis or from pilot experiments. As an interesting side product of the numerical computations of the minimum variance bound it was found that estimation of time delay influences the variances corresponding to the remaining unknown parameters unfavourably. This effect is usually overlooked in accuracy considerations found in the literature.

The procedure for synthesis of optimal test signals can only be carried out if the dynamical properties of system and noise are exactly known. In that case, however, there is no estimation problem at all. It is therefore emphasized that here the principal aim of designing optimal test signals is to obtain a reference for comparison of the performance of the usual test signals. A further aim is to derive simple rules of thumb for the selection of suitable test signals. In the latter respect this research has not yet been succesful.

In the control field the most important methods for the estimation of parameters of linear dynamic systems from additive noise corrupted responses to known inputs are the maximum likelihood (ML) method due to Åström (1965), the generalized least squares (GLS) method due to Clarke (1967) and the instrumental variable (IV) method due to Joseph, Lewis and Tou (1961). An advantage of all these methods is that the class of allowable inputs is only mildly restricted and includes spontaneous fluctuations and other statistical inputs as well. However, in most applications of the ML method and the GLS method discussed in the literature the inputs are maximum length binary sequences. The ML method and the GLS method iteratively compute both the coefficients of the discrete time system transfer function and those of the discrete time transfer function describing the noise. Furthermore, if the noise is normally distributed, the ML and GLS estimators are asymptotically efficient. A disadvantage of the ML and GLS method is that they require a relatively long computing time and a substantial amount of memory. Furthermore the iterative procedure may give rise to convergence problems. A long computation time is especially undesirable if the order and/or the time delay is not known. In that case the estimation procedure has to be repeated for increasing orders and for a number of values of time delay in order to find the best fitting solution. The importance of a short computation time for on line purposes is obvious. The IV method is computationally very simple, produces closed form solutions and requires only little computation time. On the other hand the efficiency of the IV method is hard to determine. For what follows it is important to note that the IV method requires the selection or construction of a so-called instrumental time series. The instrumental time series is correlated with causally related components of the observations of input and output, but it is independent of all non-causally related components of these observations.

Comparing the ML method and the GLS method to the one-step version of the proposed estimator, the following conclusions may be drawn:

1. The one-step procedure is computationally much simpler and is less time consuming than the ML and the GLS method. The one-step procedure requires only the Fourier analysis of the input output observations for a small number of frequencies and the solution of a set of linear equations. The one-step procedure achieves the minimum of its criterion in one single computational step. Convergence problems are therefore avoided. The ML and the GLS method are iterative. The number of computational steps is unknown in advance, while their convergence cannot always be guaranteed.
2. Using the one-step procedure a considerable data reduction can be achieved if the test signal employed either consists of a few harmonics only or has the major part of its power concentrated in a small number of dominant harmonics. The input output observations are first reduced to the Fourier coefficient estimates corresponding to the dominant harmonics. All subsequent operations are carried out on these estimates only. Binary multifrequency signals are more suitable for this procedure than sums of sinusoidal waves, since for the same harmonic content the latter signals have a larger peak factor. As opposed to the one-step procedure the ML and the GLS method operate in each step of the iteration on all points of the input output observations.
3. The one-step estimator is less accurate than the ML and the GLS method, the latter methods being efficient. Section 1.4 applies the one-step estimator to computer generated data and compares the standard deviations of the estimates so obtained to the standard deviations corresponding to the minimum variance bound. In these numerical examples the least accurate results have a

standard deviation which is about one and a half times as large as the minimum standard deviation. The most accurate results in these examples are efficient.

Next comparing the ML method and the GLS method to the proposed two-step procedure the following conclusions may be drawn:

1. The two-step procedure is still computationally simpler than the ML method and the GLS method. However, it is shown in Section 1.3 that for estimation of the variances of the residuals of the system equations, Fourier analysis of the input output observations must be carried out for a number of additional frequencies. Furthermore, in either step of the procedure a set of linear equations must be solved. These are, however, straightforward operations which require a short computation time only. Moreover, in either step of the two-step procedure convergence is achieved in ons single computational step.
2. As compared to the ML method and the GLS method the data reduction using the two-step procedure is still considerable. Due to the additional Fourier coefficients required for the estimation of the variance of the residuals the reduction is less outspoken than in the case of the one-step procedure.
3. In the numerical examples of Section 1.4 the two-step procedure achieves the minimum variance bound, as may be expected on theoretical grounds.

Finally, comparing a number of common features of the one-step and the two-step method to the properties of the ML and the GLS method, the following observations may be made:

1. The one step method and the two-step method are less general, since they are restricted to periodic inputs as opposed to the ML and the GLS method, which only require that the input is persistently exciting. A discussion of this concept will not be given here. The reader is referred to Ljung (1971), where it is shown that the class of persistently exciting inputs includes certain types of signals having continuous spectra as well as periodic inputs having sufficiently many harmonics.
2. A further difference between the one-step and the two-step estimator on one hand and the ML and the GLS method on the other is that the latter methods estimate the noise dynamics along with those of the system. The estimation of the noise dynamics is an essential part of the ML and the GLS procedure. The one-step estimator does not estimate the noise dynamics at all. The two-step estimator only estimates the spectral properties of the noise at the frequencies of the harmonics taken into account in the first step of the procedure. If, using the one-step or the two-step procedure, in addition to the system parameters the spectral properties of the noise have to be estimated, this can be done afterwards by subtracting the response of the estimated model from the observations and subsequent spectral analysis of the difference signal. Alternatively, the noise dynamics may be estimated along with those of the system by spectral analysis of the observations for frequencies corresponding to zero-valued harmonics of the test signal. Also the spectral analysis may be carried out for frequencies in between the harmonic frequencies if the observation time comprises several periods of the fundamental.
3. The ML method as well as the one-step and the two-step

procedure may be applied to systems under closed loop control. Conditions for applicability of the ML method to systems under closed loop control are discussed by Gustavsson (1974). Conditions for applicability of the GLS method in this case are not known to the author.

From a comparison of the one-step and the two-step estimator to the IV method it follows that a common characteristic of these three methods is their computational simplicity. Furthermore, the one-step method and two-step method as well as the IV method are applicable to systems under closed loop control. Using IV the test signal may be employed as instrumental time series both in the open loop and in the closed loop case. A disadvantage of the IV method is that its efficiency is hard to establish. Wong and Polak (1967) developed schemes for the construction of instrumental time series having optimal properties. Unfortunately, due to the fact that these optimal schemes are iterative, the IV method using these schemes looses much of its computational simplicity which is its most attractive property. Moreover, it is not clear how efficient these optimal schemes are, what the conditions for convergence are and whether or not these schemes can be employed if the system under investigation is under closed loop control. Furthermore, it is observed that the IV method is in general not suitable for estimation of the parameters of systems having unknown time delay, since the IV method does not employ a criterion of goodness of fit.

The estimator proposed in this thesis estimates the parameters of the differential equation of the system under test. It can, however, be shown that the procedure can easily be reformulated into an estimator of the parameters of difference equation models. In the literature virtually all schemes for estimation of parameters of linear dynamical systems apply to difference equation models. Difference equation models are extremely suitable for the digital computer, while difficulties in handling derivatives with respect to time are avoided. Moreover, from a statistical standpoint of view difference equation models may be less problematic than differential equation models. The choice of a continuous time model in this thesis is motivated as follows. First it is observed that most practical systems are continuous time. Furthermore, the proposed estimator operates on the estimates of the Fourier coefficients, not on the input output observations themselves. So frequency multiplication replaces the differentiation with respect to time. A further motivation for the choice of a continuous time model is the possibility to compute for this case an approximate expression for the covariance matrix of the estimator of both the system coefficients and the time delay. For discrete time models this problem is difficult to solve since in this case the time delay assumes discrete values only.

The outline of this thesis is as follows. Chapter 1 describes the one-step estimator and the two-step estimator and their covariance matrices. Chapter 1 also discusses numerical results of these estimators using computer generated input output observations. In order to improve the comprehensibility of Chapter 1 all proofs and mathematical details have been brought together in Chapter 3. Chapter 2 is exclusively devoted to selection and synthesis of suitable periodic inputs.

CHAPTER 1

Estimation of system parameters.

In this chapter the design of a procedure for estimating the parameters of linear, time invariant systems from noise corrupted periodic input output records is described. The parameters to be estimated are the coefficients of the differential equation and a time delay in series with the system. In what follows the coefficients of the differential equation will be referred to as system coefficients.

In the noiseless case the complex Fourier coefficient of a harmonic of a periodic input and the corresponding coefficient of the steady state response satisfy the Fourier transform of the differential equation of the system. For each harmonic this results in two linear algebraic equations in the system coefficients. These algebraic equations will be referred to as system equations. If the number of harmonics is sufficiently large an appropriate number of system equations can be selected and solved for the system coefficients.

If the observations of the periodic input and the corresponding response are corrupted by noise the Fourier coefficients of input and output can only be estimated. Generally these estimates do not satisfy the system equations. The estimator for the system coefficients proposed in this research minimizes the residuals of any number of system equations in a weighted least squares sense. This estimator is discussed in Section 1.1. With respect to the system coefficients the estimator is a closed form expression. The time delay is estimated by repeating the procedure for a number of values of time delay and selecting the optimal solution. The estimator is consistent if the following conditions are both satisfied: 1) the estimator of the Fourier coefficients is consistent and 2) the total number of unknown parameters is smaller than or equal to twice the number of harmonics taken into consideration. Since these conditions do not require input and output noises to be independent, the estimator is also suitable for the case that the system is under closed loop control.

In Section 1.2 the asymptotic covariance matrix of the estimator is discussed. From this covariance matrix it follows how the weights of the least squares estimator must be chosen in order to minimize the variance. Section 1.3 describes a two-step procedure for estimating these optimum weights along with the parameters.

Finally, in Section 1.4 some numerical results computed from computer generated data are described.

In order to improve the readability of this chapter proofs and mathematical details have been omitted. These are discussed in Chapter 3.

1.1 The least squares estimator.

The applicability of the parameter estimation procedure proposed in this research is restricted to time-invariant systems describable by the following scalar differential difference equation

$$a_0 y(t) + a_1 \frac{dy(t)}{dt} + \ldots + a_N \frac{d^N y(t)}{dt^N} =$$

$$= b_o\, u(t - \tau) + b_1 \frac{du(t - \tau)}{dt} + \ldots +$$

$$+ b_M \frac{d^M u(t - \tau)}{dt^M} \tag{1.1.1}$$

where u(t) is the input signal, y(t) is the response to u(t) and τ denotes time delay. For simplicity the arguments in (1.1.1) will be referred to as time. However, these arguments may denote some other variable such as distance. Furthermore it is assumed that the system is stable and that $b_M = 1$.

Now let u(t) be a periodic test signal with period T and let y(t) be the steady state response to u(t). Define the complex Fourier coefficient of the kth harmonic of u(t) by

$$\gamma_{ku} = \alpha_{ku} - j\beta_{ku} = \frac{1}{T} \int_o^T u(t) \exp(-j2\pi kt/T)\, dt \tag{1.1.2}$$

and define γ_{ky}, α_{ky} and β_{ky} correspondingly. For the moment let τ be known and for simplicity assume $\tau = 0$. Then

$$(a_0 + a_1 s_k + \ldots + a_N s_k{}^N)\gamma_{ky} =$$

$$=(b_0 + b_1 s_k + \ldots + s_k{}^M)\gamma_{ku} \tag{1.1.3}$$

where $s_k = j2\pi k/T$. The complex equation (1.1.3) is equivalent to the real equations.

$$\mathrm{Re}\{(a_0 + a_1 s_k + \ldots + a_N s_k{}^N)\gamma_{ky} +$$

$$-(b_0 + b_1 s_k + \ldots + b_{M-1} s_k{}^{M-1})\gamma_{ku}\} =$$

$$= \mathrm{Re}(s_k^{\,M}\gamma_{ku}) \tag{1.1.4}$$

and

$$\mathrm{Im}\{(a_0 + a_1 s_k + \ldots + a_N s_k^{\,N})\gamma_{ky} +$$
$$- (b_0 + b_1 s_k + \ldots + b_{M-1} s_k^{\,M-1})\gamma_{ku}\} =$$
$$= \mathrm{Im}(s_k^{\,M}\gamma_{ku}) \tag{1.1.5}$$

These are two simultaneous, inhomogeneous, linear algebraic equations in the system coefficients. The coefficient matrix of these equations is

$$P_k = \begin{pmatrix} \mathrm{Re}\ s_k^0\ \gamma_{ky} & \mathrm{Re}\ s_k^1\ \gamma_{ky} & \ldots & \mathrm{Re}\ s_k^N\ \gamma_{ky} & - \mathrm{Re}\ s_k^0\ \gamma_{ku} & - \mathrm{Re}\ s_k^1\ \gamma_{ku} & \ldots & - \mathrm{Re}\ s_k^{M-1}\ \gamma_{ku} \\ \mathrm{Im}\ s_k^0\ \gamma_{ky} & \mathrm{Im}\ s_k^1\ \gamma_{ky} & \ldots & \mathrm{Im}\ s_k^N\ \gamma_{ky} & - \mathrm{Im}\ s_k^0\ \gamma_{ku} & - \mathrm{Im}\ s_k^1\ \gamma_{ku} & \ldots & - \mathrm{Im}\ s_k^{M-1}\ \gamma_{ku} \end{pmatrix} \tag{1.1.6}$$

Consequently if L harmonics of u(t) and y(t) with harmonic numbers $k_1, \ldots, k_L$ are taken into consideration, 2L equations result with coefficient matrix

$$P = (P'_{k_1}\ P'_{k_2}\ \ldots\ P'_{k_L})' \tag{1.1.7}$$

where the prime denotes transposition. Defining

$$d = (\mathrm{Re}\ s_{k_1}^M \gamma_{k_1 u}\ \mathrm{Im}\ s_{k_1}^M \gamma_{k_1 u}\ \ldots .\ \mathrm{Re}\ s_{k_L}^M \gamma_{k_L u}$$
$$\mathrm{Im}\ s_{k_L}^M \gamma_{k_L u})' \tag{1.1.8}$$

and

$$c = (a_0\ a_1\ \ldots ..\ a_N\ b_0\ \ldots .\ b_{M-1})' \tag{1.1.9}$$

the 2L equations may thus be written in vector notation

$$Pc = d \tag{1.1.10}$$

Now first suppose that u(t) and y(t) can exactly be measured. Then $\alpha_{k_i y}$, $\beta_{k_i y}$, $\alpha_{k_i u}$ and $\beta_{k_i u}$ can exactly be computed. Consequently the elements of P and d in (1.1.10) are known. If it is assumed that $N+M+1 \leqslant 2L$, $N+M+1$ equations can be selected from (1.1.10) and solved for c.

In the noiseless case with unknown time delay the procedure is as follows. Assume that $2L \geqslant N+M+2$. First $\gamma_{k_i u} \exp -s_k \tau$ is substituted for all $\gamma_{k_i u}$ in (1.1.10). Next N+M+1 equations are selected from (1.1.10) and solved for for c for a number of selected values of τ. The value of τ and the corresponding solution for c which satisfy the remaining equation(s) of (1.1.10) are selected.

From the above considerations it is clear that the parameters

$$\theta = (a_0 \ \ldots \ a_N \ b_0 \ \ldots \ b_{M-1} \ \tau)' \qquad (1.1.11)$$

of the model (1.1.1) can easily be computed if it is assumed that u(t) and y(t) can exactly be measured. In practice this assumption is unrealistic. For example, under normal operating conditions responses to additional system inputs are superimposed on the response to the test signal. Furthermore the observations made on the response may be corrupted by measurement errors or internal system noise.

If the system is under closed loop control the test signal is usually introduced at the set point. Consequently the input of the system is not the test signal itself and has to be measured. Therefore in this case the observations of both input and output are subject to disturbances. The same situation occurs if the test signal is applied to the system using a transducer with unknown dynamic characteristics.

In what follows it will be assumed that the disturbances are adequately described by additive stationary stochastic processes. So the observations at the output and input are described by

$$z(t) = y(t) + h(t) \tag{1.1.12}$$

and

$$v(t) = u(t) + g(t) \tag{1.1.13}$$

respectively, where g(t) and h(t) are stationary stochastic processes. Furthermore it is assumed that the observation time is an integral number J of periods of the input signal u(t). Now, define

$$\hat{\gamma}_{ku}^{(J)} = \hat{\alpha}_{ku}^{(J)} - j\hat{\beta}_{ku}^{(J)} = \frac{1}{JT}\int_0^{JT} v(t)\exp(-j2\pi kt/T)dt \tag{1.1.14}$$

and define $\hat{\gamma}_{ky}^{(J)}$, $\hat{\alpha}_{ky}^{(J)}$ and $\hat{\beta}_{ky}^{(J)}$ correspondingly. In Section 3.1 it is shown that $\hat{\gamma}_{ku}^{(J)}$ and $\hat{\gamma}_{ky}^{(J)}$ are consistent least sqares estimators of γ_{ku} and γ_{ky} respectively if

$$\int_{-\infty}^{+\infty} |R_{gg}(t)|dt < \infty \text{ and } \int_{-\infty}^{+\infty} |R_{hh}(t)|dt < \infty \tag{1.1.15}$$

where $R_{gg}(\tau)$ and $R_{hh}(\tau)$ denote the autocovariance functions of g(t) and h(t) respectively. The condition (1.1.15) is a sufficient condition, not a necessary one. For example, (1.1.15) is not satisfied if g(t) and h(t) have periodic components. Nevertheless it is shown in Section 3.1 that $\hat{\gamma}_{ky}^{(J)}$ and $\hat{\gamma}_{ku}^{(J)}$ are consistent in this case provided that the frequencies of the periodic components do not coincide with $2\pi k/T$.

It is important to note that the stationarity assumption requires that trends in the mean of v(t) and z(t) are removed. Schemes for the elimination of trends will be discussed in the last part of this section.

Now, consider equation (1.1.10). Let $\hat{P}^{(J)}$ be the matrix which is obtained by substituting $\hat{\gamma}^{(J)}_{k_i u}$ for all $\gamma_{k_i u}$ and $\hat{\gamma}^{(J)}_{k_i y}$ for all $\gamma_{k_i y}$ respectively in P. Define the vector $\hat{d}^{(J)}$ correspondingly. Then generally

$$\hat{P}^{(J)}c - \hat{d}^{(J)} \neq 0$$

The vector

$$\hat{\varepsilon}^{(J)} = \hat{P}^{(J)}c - \hat{d}^{(J)} \tag{1.1.16}$$

will be referred to as the vector of residuals. From (1.1.4) - (1.1.9) it follows that the elements of $\hat{\varepsilon}^{(J)}$ are of the form

$$\mathrm{Re}\{A(j\omega_{k_i})\ \hat{\gamma}^{(J)}_{k_i y} - B(j\omega_{k_i})\exp(-j\omega_{k_i}\tau)\hat{\gamma}^{(J)}_{k_i u}\} \tag{1.1.17}$$

and

$$\mathrm{Im}\{A(j\omega_{k_i})\hat{\gamma}^{(J)}_{k_i y} - B(j\omega_{k_i})\exp(-j\omega_{k_i}\tau)\hat{\gamma}^{(J)}_{k_i u}\} \tag{1.1.18}$$

Furthermore let Ω^{-1} be a symmetric, positive definite matrix. Then

$$\hat{\varepsilon}'^{(J)}\ \Omega^{-1}\ \hat{\varepsilon}^{(J)} \tag{1.1.19}$$

is a positive quadratic form in the residuals. In this research the estimator of the coefficients is taken as that solution $\hat{c}^{(J)}$ for c which minimizes (1.1.19). The matrix Ω^{-1} is included in this expression to allow for weighting schemes which take a priori knowledge about system and noise into account. It is shown in Section 3.2 that $\hat{c}^{(J)}$ satisfies

$$\hat{c}^{(J)} = (\hat{P}'^{(J)}\ \Omega^{-1}\hat{P}^{(J)})^{-1}\ \hat{P}'^{(J)}\ \Omega^{-1}\hat{d}^{(J)} \tag{1.1.20}$$

In what follows this estimator will be referred to as weighted least squares estimator or simply least squares estimator of the system coefficients. If no a priori knowledge about the system and noise characteristics is available Ω^{-1} will be taken as the identity matrix. A more refined choice is discussed in Section 1.3.

Consistency of $\hat{c}^{(J)}$.

It is shown in Section 3.2 that $\hat{c}^{(J)}$ is a consistent estimator of c if the following conditions are both satisfied:

1) the estimator of the Fourier coefficients is consistent and
2) the number of unknown system coefficients is smaller than or equal to twice the number of non-zero harmonics taken into consideration.

It should be noted that these conditions are very general. For instance, no assumptions are made with respect to the amplitude distributions of g(t) and h(t). Furthermore the inequalities (1.1.15) constitute the only restrictions on the class of allowable power density spectra of g(t) and h(t).

Estimation of the order.

In order to determine the order N of the left-hand member and the order M of the right-hand member of (1.1.1), (1.1.20) is solved for increasing M and N. The optimal M and N and the corresponding solution for $\hat{c}^{(J)}$ are selected.

Estimation of time delay.

Time delay is estimated by substituting $\hat{\gamma}_{k_i u} \exp - s_{k_i} \tau$ for all $\hat{\gamma}^{(J)}_{k_i u}$ in (1.1.20) and solving (1.1.20) for $\hat{c}^{(J)}$ for a number of selected values of τ. The value $\hat{\tau}^{(J)}$ of τ and the corresponding solution for $\hat{c}^{(J)}$ which minimize the criterion (1.1.19) are selected. In Section 3.2 it is shown that this procedure estimates the parameters θ consistently if the following conditions are both satisfied: 1) the estimator of the Fourier coefficients is consistent and 2) the number of unknown parameters is smaller than or equal to twice the number of non-zero harmonics taken into consideration.

In the case of all-pole models the elements of the matrix $\hat{P}^{(J)}$ in (1.1.20) are independent of the Fourier coefficients $\hat{\gamma}^{(J)}_{k_i u}$. Therefore substitution of $\hat{\gamma}_{k_i u} \exp - s_{k_i} \tau$ for $\hat{\gamma}^{(J)}_{k_i u}$ in (1.1.20) only changes the vector $\hat{d}^{(J)}$. So if (1.1.20) is solved for $\hat{c}^{(J)}$ for a number of values of τ, the matrix $(\hat{P}'^{(J)}\Omega^{-1}\hat{P}^{(J)})^{-1}\,\hat{P}'^{(J)}\Omega^{-1}$ has to be computed only once. It should be noted that as a result of the periodicity of the test signal the time delay can only be estimated modulo the period of the fundamental.

Estimation in closed loop.

If the system is under closed loop control the noise at the input and the noise at the output are correlated through the feedback path. As a result a number of conventional open loop estimation schemes are generally no longer consistent in the closed loop situation. For a discussion the reader is referred to Gustavsson, Ljung and Söderström (1974). Irrespective as to whether the noises at the input and output are correlated or not, the estimator proposed in this research is consistent if the estimator of the Fourier coefficients is consistent and if, in addition, the number of harmonics taken into consideration is sufficiently large. Therefore this estimator is equally well applicable to open loop systems as to systems under closed loop control.

Computational aspects.

The numerical procedure for computation of the estimates involves Fourier analysis of v(t) and z(t) and subsequent solution of the set of linear equations (1.1.20). In the case of unknown time delay or unknown order of the system only the second step of this procedure has to be repeated for a number of values of time delay or for various orders in order to find the best fit. The Fourier analysis needs to be carried out only once.

The Fourier analysis may be carried out directly or by means of the fast Fourier transform (FFT). The FFT is a relatively fast and accurate algorithm for computing discrete Fourier transforms (DFT). For a discussion of the FFT see Gentleman (1966). Now suppose that a record consists of n samples, where n is a power of two. The FFT requires for the DFT of the record $n \log_2 n$ complex multiplications, whereas a direct approach requires n^2 complex multiplications. However, it should be noted that using the direct approach the DFT can be computed for each harmonic individually whereas the FFT computes the DFT for all harmonics simultaneously. So if L harmonics are taken into consideration the ratio of the number of multiplications using the direct approach to the number of multiplications required by the FFT is given by $L/\log_2 n$.

The requirement that the number of unknown parameters must be smaller than or equal to twice the number of harmonics taken into consideration constitutes the lower bound on L. It is clear that for systems of low order L may be quite small. For example for a second order all-pole model with time delay L = 2 is sufficient. So in this case the direct approach will be much faster since it is reasonable to assume that n >> 4.

Suitable test signals for testing at a limited number of frequencies are binary multifrequency test signals. These are periodic two-level signals that have the major part of their power concentrated in a limited number of relatively widely spaced harmonics. The construction of binary multifrequency signals is discussed in Section 2.3 and Section 2.4. A number of computer experiments using binary multifrequency test signals is described in Section 1.4.

Removal of trends.

In practice the input observations v(t) and the output observations z(t) may exhibit trends in their mean values. This means that the mean value of the input noise g(t) and that of the output noise h(t) are nonstationary. In Section 3.1 the unbiasedness and mean square convergence of the estimators of the Fourier coefficients are proved under stationarity assumptions. In order to illustrate the effect of trends on the estimators of the Fourier coefficients consider the following example. Let the mean of h(t) be a trend described by the second degree polynomial

$$E\,h(t) = \rho_1 t + \rho_2 t^2$$

where ρ_1 and ρ_2 are constant coefficients. Then it is straightforward to show that

$$E[\hat{\alpha}_{ky}^{(J)}] - \alpha_{ky} = \rho_2 \frac{2}{\omega_k^2}$$

and

$$E[\hat{\beta}_{ky}^{(J)}] - \beta_{ky} = -\rho_1 \frac{1}{\omega_k} - \rho_2 \frac{JT}{\omega_k}$$

From these expressions it is clear that as a result of the trend $\hat{\alpha}_{ky}^{(J)}$ and $\hat{\beta}_{ky}^{(J)}$ may seriously be biased. So the question is how to remove the trend. This problem can be solved employing techniques used in econometry for decomposing a time series into a trend and a periodic seasonal component. From the literature two major approaches for this decomposition are available. These are polynomial curve fitting and moving average procedures. See Wonnacott and Wonnacott (1970), page 68.

Polynomial curve fitting is not very suitable for the detrending of signals of the type considered in this thesis. The reason is that both the Fourier coefficients of the periodic component and the coefficients of the trend polynomial have to be estimated simultaneously in the same least squares procedure. Estimation of the polynomial coefficients prior to estimation of the Fourier coefficients gives rise to wrong results. The simultaneous estimation procedure requires the solving of a set of linear equations both for all Fourier coefficients of the periodic component and the coefficients of the trend polynomial. The computation of the coefficients of this set of equations and the subsequent solution constitute a substantial computational effort and become prohibitive with many harmonics.

The moving average procedure for removal or reduction of the trend is computationally much simpler. This procedure may be described as follows. Let again $z(t)$ be the observations to be detrended. Then $z(t)$ is transformed into $\bar{z}(t)$ as follows

$$\bar{z}(t) = z(t) - \frac{1}{T} \int_{t - \frac{1}{2} T}^{t + \frac{1}{2} T} z(t_1) dt_1$$

Since the average of $y(t)$ over the fundamental period T is zero this expression may be written

$$\bar{z}(t) = y(t) + h(t) - \frac{1}{T} \int_{t - \frac{1}{2}T}^{t + \frac{1}{2}T} h(t_1)dt_1$$

Note that with this transformation of $z(t)$ into $\bar{z}(t)$ the first and the last half period of the observations are lost. Also note that this procedure leaves the amplitude and phase of the harmonics of $y(t)$ unchanged. Now let again

$$E\, h(t) = \rho_1 t + \rho_2 t^2$$

Then

$$E[h(t) - \frac{1}{T}\int_{t - \frac{1}{2}T}^{t + \frac{1}{2}T} h(t_1)dt_1] = \rho_2 \frac{T^2}{12}$$

and hence

$$E\, \bar{z}(t) = y(t) + \text{constant}$$

Therefore, if the Fourier coefficients of $y(t)$ are estimated from $\bar{z}(t)$ instead of from $z(t)$, it follows that

$$E\hat{\alpha}_{ky}^{(J)} = \alpha_{ky}$$

and

$$E\hat{\beta}_{ky}^{(J)} = \beta_{ky}$$

This example shows that by simply subtracting the moving average the effect of linear and quadratic trend terms is completely eliminated.

1.2. The asymptotic covariance matrix of the least squares estimator.

This section discusses the asymptotic covariance matrix of the estimator $\hat{\theta}^{(J)} = (\hat{a}_0^{(J)}, \hat{a}_1^{(J)}, \ldots, \hat{a}_N^{(J)}, \hat{b}_0^{(J)}, \ldots, \hat{b}_{M-1}^{(J)}, \hat{\tau}^{(J)})'$. Recall that the superscript J in this expression refers to the observation time JT. For finite J the expression for the covariance matrix of $\hat{\theta}^{(J)}$ is very difficult to obtain. However, it will be shown in Section 3.4 that the computation of the asymptotic covariance matrix of $\hat{\theta}^{(J)}$ is straightforward. The asymptotic covariance matrix of $\hat{\theta}^{(J)}$ is defined as follows. Let

$$V_{\hat{\theta}\hat{\theta}} = \lim_{J \to \infty} E\{\sqrt{JT}(\hat{\theta}^{(J)} - E\hat{\theta}^{(J)})\}\{\sqrt{JT}(\hat{\theta}^{(J)} - E\hat{\theta}^{(J)})\}'$$

and assume that $V_{\hat{\theta}\hat{\theta}}$ is a matrix of finite constants. Then the asymptotic covariance matrix $\Sigma_{\hat{\theta}\hat{\theta}}$ of $\hat{\theta}^{(J)}$ is defined by

$$\Sigma_{\hat{\theta}\hat{\theta}} = \frac{1}{JT} V_{\hat{\theta}\hat{\theta}}$$

This asymptotic covariance matrix will be used as an approximation to the covariance matrix for finite J.

First a more specific model for the noise in the observations of the input and in those of the output is set up. The noise in the observations of the input is assumed to be composed as follows

$$g(t) = g_1(t) + g_2(t) \tag{1.2.1}$$

where $g_1(t)$ and $g_2(t)$ are stationary stochastic processes. The process $g_2(t)$ represents measurement noise. For example $g_2(t)$ may be quantization noise. The process $g_1(t)$ is an additional input to the system and may represent the normal operating input signal.

The noise in the observations of the output is described by

$$h(t) = h_1(t) + h_2(t) + h_3(t) \tag{1.2.2}$$

where $h_1(t)$, $h_2(t)$ and $h_3(t)$ are stationary stochastic processes. The process $h_1(t)$ is the steady state response to the process $g_1(t)$ and is therefore correlated with $g_1(t)$. The process $h_3(t)$ is an equivalent disturbance at the output representing internal noises in the system. In order to include the case that the system is under closed loop control it is assumed that the processes $g_1(t)$, $h_1(t)$ and $h_3(t)$ are mutually correlated. The process $h_2(t)$ represents measurement noise in the observations of the output and is assumed to be independent of all other processes. The same assumption is made with respect to $g_2(t)$. Finally, it is assumed that $g(t)$ and $h(t)$ are independent of $u(t)$ and $y(t)$.

In what follows it is supposed that the conditions for consistency of $\hat{\theta}^{(J)}$ are satisfied and that in addition

$$\lim_{JT\to\infty} \int_{-JT}^{JT} (1 - |t|/JT) R_{gg}(t) \exp(-j2\pi kt/T)dt = \\ = S_{gg}(j2\pi k/T) \qquad (1.2.3)$$

$$\lim_{JT\to\infty} \int_{-JT}^{JT} (1 - |t|/JT) R_{hh}(t) \exp(-j2\pi kt/T)dt = \\ = S_{hh}(j2\pi k/T) \qquad (1.2.4)$$

and

$$\lim_{JT\to\infty} \int_{-JT}^{JT} (1 - |t|/JT) R_{gh}(t) \exp(-j2\pi kt/T)dt = \\ = S_{gh}(j2\pi k/T) \qquad (1.2.5)$$

where $S_{gg}(j\omega)$ and $S_{hh}(j\omega)$ denote the power density spectra of $g(t)$ and $h(t)$ respectively and $S_{gh}(j\omega)$ and $R_{gh}(\tau)$ are the cross power density spectrum and cross covariance function of these processes respectively. The assumptions (1.2.3) - (1.2.5) are discussed in Section 3.1 and Section 3.3. Furthermore it is assumed that the higher-than-two order central moments of the estimator of the Fourier coefficients are of order of magnitude lower than $1/JT$. Now, define

$$\Phi = (\Phi'_{k_1}\ \Phi'_{k_2}\ \dots\ \Phi'_{k_i}\ \dots\ \Phi'_{k_L})' \tag{1.2.6}$$

where

$$\Phi_{k_i} = \begin{pmatrix} \mathrm{Re}\ s^0_{k_i}\gamma_{k_i y} & \mathrm{Re}\ s^1_{k_i}\gamma_{k_i y} & \dots & \mathrm{Re}\ s^N_{k_i}\gamma_{k_i y} & -\mathrm{Re}\ s^0_{k_i}\gamma_{k_i u}\exp(-s_{k_i}\tau) & -\mathrm{Re}\ s^1_{k_i}\gamma_{k_i u}\exp(-s_{k_i}\tau) & \dots & -\mathrm{Re}\ s^{M-1}_{k_i}\gamma_{k_i u}\exp(-s_{k_i}\tau) & -\mathrm{Re} - s_{k_i}B(s_{k_i})\exp(-s_{k_i}\tau) \\ \mathrm{Im}\ s^0_{k_i}\gamma_{k_i y} & \mathrm{Im}\ s^1_{k_i}\gamma_{k_i y} & \dots & \mathrm{Im}\ s^N_{k_i}\gamma_{k_i y} & -\mathrm{Im}\ s^0_{k_i}\gamma_{k_i u}\exp(-s_{k_i}\tau) & -\mathrm{Im}\ s^1_{k_i}\gamma_{k_i u}\exp(-s_{k_i}\tau) & \dots & -\mathrm{Im}\ s^{M-1}_{k_i}\gamma_{k_i u}\exp(-s_{k_i}\tau) & -\mathrm{Im} - s_{k_i}B(s_{k_i})\exp(-s_{k_i}\tau) \end{pmatrix} \tag{1.2.7}$$

where $s_{k_i} = j2\pi k_i/T$. Furthermore let

$$\Lambda = \mathrm{diag}(\lambda(\omega_{k_1}),\ \lambda(\omega_{k_1}),\ \dots,\ \lambda(\omega_{k_i}),\ \lambda(\omega_{k_i}),\ \dots,\ \lambda(\omega_{k_L}),\ \lambda(\omega_{k_L})) \tag{1.2.8}$$

where

$$\lambda(\omega_{k_i}) = \tfrac{1}{2}[\,|A(s_{k_i})|^2\{S_{h_2h_2}(s_{k_i}) + S_{h_3h_3}(s_{k_i})\} + |B(s_{k_i})|^2 S_{g_2g_2}(s_{k_i})] \tag{1.2.9}$$

Then it is shown in Section 3.4 that the asymptotic covariance matrix $\Sigma_{\hat\theta\hat\theta}$ of $\hat\theta^{(J)}$ may be written

$$\Sigma_{\hat\theta\hat\theta} = \frac{1}{JT}\,(\Phi'\Omega^{-1}\Phi)^{-1}\Phi'\Omega^{-1}\Lambda\Omega^{-1}\Phi(\Phi'\Omega^{-1}\Phi)^{-1} \tag{1.2.10}$$

and that

$$\Sigma_{\hat{\varepsilon}\hat{\varepsilon}} = \frac{1}{JT} \Lambda \qquad (1.2.11)$$

represents the asymptotic covariance matrix of the residuals.

Discussion.

The equations (1.2.8), (1.2.9) and (1.2.10) show that the processes $g_1(t)$ and $h_1(t)$ do not contribute to the elements of $\Sigma_{\hat{\theta}\hat{\theta}}$. Intuitively this may be explained as follows. First it is observed that $g_1(t)$ and $h_1(t)$ satisfy the differential equation of the system, since $h_1(t)$ is the response to $g_1(t)$. The contribution of $g_1(t)$ to the error in $\hat{\gamma}_{ku}^{(J)}$ and the contribution of $h_1(t)$ to the error in $\hat{\gamma}_{ky}^{(J)}$ are finite Fourier transforms of $g_1(t)$ and $h_1(t)$ respectively. These finite Fourier transforms approximately satisfy the equations $Pc - d = 0$. Clearly this results in a contribution to the elements of the covariance matrix of the residuals, and therefore to the elements of the covariance matrix of $\hat{\theta}^{(J)}$, of order lower than J^{-1}. According to the definition of the asymptotic covariance matrix contributions of order lower than J^{-1} are neglected. So the fact that $\Sigma_{\hat{\theta}\hat{\theta}} = 0$ for $h_2(t) = h_3(t) = g_2(t) = 0$ only indicates that in this particular case the elements of the asymptotic covariance matrix are of order lower than J^{-1}.

The expression for $\Sigma_{\hat{\theta}\hat{\theta}}$ is similar to the expression for the covariance matrix of the estimator of the coefficients of the generalized linear regression model. See Eykhoff (1974), Section 6.1. Now let $\tilde{\theta}^{(J)}$ be the weighted least squares estimator for the case that $\Omega = \Lambda$. It is known from regression theory that $\Sigma_{\tilde{\theta}\tilde{\theta}}$ is smaller than $\Sigma_{\hat{\theta}\hat{\theta}}$ for any other choice of Ω. See Goldberger (1964), page 233. From (1.2.10) it follows that the asymptotic covariance matrix

$\Sigma_{\tilde{\theta}\tilde{\theta}}$ of $\tilde{\theta}^{(J)}$ may be written

$$\Sigma_{\tilde{\theta}\tilde{\theta}} = \frac{1}{JT} (\Phi'\Lambda^{-1}\Phi)^{-1} \tag{1.2.12}$$

The estimator $\tilde{\theta}^{(J)}$ will be referred to as minimum variance least squares estimator.

1.3. Two-step least squares estimation of the system parameters.

It is clear that the minimum variance property of the estimator $\tilde{\theta}^{(J)}$ described in the preceding section is a desirable one. In practice, however, Λ is not known since its computation requires detailed knowledge about the dynamic characteristics of system and noise. In this section a two-step algorithm is described for estimating the elements of Λ along with the parameters. The aim is to obtain an estimator whose properties are comparable to those of the minimum variance estimator discussed in the preceding section.

Define

$$V(k) = \frac{1}{JT} \int_0^{JT} v(t)\exp(-j2\pi kt/JT)dt \tag{1.3.1}$$

and

$$Z(k) = \frac{1}{JT} \int_0^{JT} z(t)\exp(-j2\pi k/JT)dt \tag{1.3.2}$$

respectively. In the first step of the procedure V(k) and Z(k) are computed for $k = k_iJ - k_w, \ldots, k_iJ-1, k_iJ, k_iJ+1, \ldots, k_iJ + k_w$, where $i = 1, \ldots, L$ and k_w is a fixed integer. Note that $\hat{\gamma}^{(J)}_{k_i u} = V(k_iJ)$ and $\hat{\gamma}^{(J)}_{k_i y} = Z(k_iJ)$. Next the parameters $\hat{\theta}^{(J)}$ are estimated from the Fourier coefficient estimates corresponding to the harmonic numbers $k_1, \ldots, k_L$ using the least-squares estimator

$$\hat{c}^{(J)} = (\hat{P}'^{(J)}\hat{P}^{(J)})^{-1} \hat{P}'^{(J)}\hat{d}^{(J)} \tag{1.3.3}$$

This is the weighted least squares estimator with weighting matrix $\Omega^{-1} = I$, where I is the identity matrix. In the case of unknown time delay the parameters are estimated by computing $\hat{c}^{(J)}$ for a number of selected values of τ and selecting the value $\hat{\tau}^{(J)}$ and the corresponding solution for $\hat{c}^{(J)}$ which minimize

$$(\hat{P}^{(J)}c - \hat{d}^{(J)})' (\hat{P}^{(J)}c - \hat{d}^{(J)}) \tag{1.3.4}$$

These estimates are denoted by $\hat{\theta}_I^{(J)}$ where the subscript I refers to a weighting matrix $\Omega^{-1} = I$.

In the second step first the elements $\lambda(\omega_{k_i})$ of Λ are estimated. According to (1.2.8) Λ is diagonal. Furthermore it follows from (1.2.11) that

$$\Lambda = JT\ \Sigma_{\hat{\varepsilon}\hat{\varepsilon}}$$

where $\Sigma_{\hat{\varepsilon}\hat{\varepsilon}}$ is the asymptotic covariance matrix of the residuals $\hat{\varepsilon}^{(J)}$. According to (1.1.17) and (1.1.18) the elements of $\hat{\varepsilon}^{(J)}$ are of the form

$$\mathrm{Re}\{A(j\omega_{k_i})\hat{\gamma}_{k_i y}^{(J)} - B(j\omega_{k_i})\exp(-j\omega_{k_i}\tau)\hat{\gamma}_{k_i u}^{(J)}\} \tag{1.3.5}$$

and

$$\mathrm{Im}\{A(j\omega_{k_i})\hat{\gamma}_{k_i y}^{(J)} - B(j\omega_{k_i})\exp(-j\omega_{k_i}\tau)\hat{\gamma}_{k_i u}^{(J)}\} \tag{1.3.6}$$

The expectation of the residuals (1.3.5) and (1.3.6) is zero since $\hat{\gamma}_{k_i y}^{(J)}$ and $\hat{\gamma}_{k_i u}^{(J)}$ are unbiased estimators of the Fourier coefficients $\gamma_{k_i y}$ and $\gamma_{k_i u}$ respectively. Furthermore, according to (1.2.8), (1.2.9) and (1.2.11) the asymptotic variance of the residual (1.3.5) and that of the residual (1.3.6) are both equal to

$$\frac{1}{JT}\lambda(\omega_{k_i}) = \frac{1}{2JT}\left[\,|A(j\omega_{k_i})|^2\{S_{h_2h_2}(j\omega_{k_i}) + S_{h_3h_3}(j\omega_{k_i})\}\right.$$

$$\left. + |B(j\omega_{k_i})|^2\, S_{g_2g_2}(j\omega_{k_i})\right] \tag{1.3.7}$$

The residuals (1.3.5) and (1.3.6) can not exactly be computed from the estimated Fourier coefficients $\hat{\gamma}_{k_i y}^{(J)}$ and $\hat{\gamma}_{k_i u}^{(J)}$ since the system parameters are not known. However, using the system parameter estimates obtained in the first step it is proposed here to estimate these residuals by means of the estimators

$$\mathrm{Re}\{A(j\omega_{k_i})\hat{\gamma}^{(J)}_{k_i y} - B(j\omega_{k_i})\exp(-j\omega_{k_i}\tau)\ \hat{\gamma}^{(J)}_{k_i u}\}_{\theta\ =\ \hat{\theta}^{(J)}_I}$$

and

$$\mathrm{Im}\{A(j\omega_{k_i})\hat{\gamma}^{(J)}_{k_i y} - B(j\omega_{k_i})\exp(-j\omega_{k_i}\tau)\ \hat{\gamma}^{(J)}_{k_i u}\}_{\theta\ =\ \hat{\theta}^{(J)}_I}$$

respectively. These quantities will be referred to as the measured residuals. Then one might think of estimating the variance of the true residuals from the measured residuals using the following estimator

$$\tfrac{1}{2}[\mathrm{Re}^2\{A(j\omega_{k_i})\hat{\gamma}^{(J)}_{k_i y} - B(j\omega_{k_i})\exp(-j\omega_{k_i}\tau)\hat{\gamma}^{(J)}_{k_i u}\}_{\theta\ =\ \hat{\theta}^{(J)}_I} +$$

$$+ \mathrm{Im}^2\{A(j\omega_{k_i})\hat{\gamma}^{(J)}_{k_i y} - B(j\omega_{k_i})\exp(-j\omega_{k_i}\tau)\hat{\gamma}^{(J)}_{k_i u}\}_{\theta\ =\ \hat{\theta}^{(J)}_I}] \tag{1.3.8}$$

However, selecting (1.3.8) as an estimator of the variance of the residual (1.3.5) and that of the residual (1.3.6) is equivalent to estimating the variance of a zero mean random variable from only two observations on that variable. In that case the standard deviation and the expectation of the estimator are of the same order of magnitude. So it is concluded that (1.3.8) is a very inaccurate estimator. In view of (1.3.7) the quantities to be estimated, that is the asymptotic variances of the residuals, are power spectral density functions. In spectral analysis the procedure for reducing the variance of a spectral estimator is to introduce a window. See Blackman and Tukey (1958). This implies that the spectral estimator for a particular frequency is replaced by a weighted average of estimators over neighbouring frequencies, the weights being determined by the spectral window. For reducing the variance of the estimator (1.3.8) of the asymptotic variance of the residuals, this estimator is changed into the uniformly weighted sum $\frac{1}{JT}\hat{\lambda}_w(\omega_{k_i})$ defined by

$$\frac{1}{JT}\hat{\lambda}_w(\omega_{k_i}) = \frac{1}{2(2k_w+1)} \sum_{k=k_iJ-k_w}^{k_iJ+k_w} \mathrm{Re}^2\{A(j2\pi k/JT)Z(k) +$$

$$-B(j2\pi k/JT)\exp(-j2\pi k\tau/JT)V(k)\}_{\theta = \hat{\theta}_I^{(J)}} +$$

$$+\mathrm{Im}^2\{A(j2\pi k/JT)Z(k) +$$

$$-B(j2\pi k/JT)\exp(-j2\pi k\tau/JT)V(k)\}_{\theta = \hat{\theta}_I^{(J)}} \qquad (1.3.9)$$

where $i = 1, \ldots, L$. Note that the estimates $\hat{\lambda}_w(\omega_{k_i})$ can be computed from the results of the first step of the procedure, that is from $\hat{\theta}_I^{(J)}$ and $V(k)$, $Z(k)$ for $k = k_iJ - k_w, \ldots, k_iJ + k_w$ and $i = 1, \ldots, L$. According to the theory of spectral analysis k_w must be small in order to keep the bias of $\hat{\lambda}_w(\omega_{k_i})$ small. On the other hand in order to keep the variance small, k_w must be large. Roughly speaking the ratio of the variance of $\hat{\lambda}_w(\omega_{k_i})$ to the square of $\lambda(\omega_{k_i})$ is $1/(4k_w + 2)$. So one is forced to compromise between the variance and the bias of the estimator $\hat{\lambda}_w(\omega_{k_i})$. For a detailed discussion of spectral windows see Jenkins and Watts (1967).

From the estimator $\hat{\lambda}_w(\omega_{k_i})$ the estimator $\hat{\Lambda}$ of the matrix Λ is constructed as follows

$$\hat{\Lambda} = \mathrm{diag}\,(\hat{\lambda}_w(\omega_{k_1}), \hat{\lambda}_w(\omega_{k_1}), \ldots, \hat{\lambda}_w(\omega_{K_L}), \hat{\lambda}_w(\omega_{K_L})) \qquad (1.3.10)$$

Next the solution $\hat{\hat{\theta}}^{(J)}$ for θ is computed which minimizes

$$(\hat{P}^{(J)}c - \hat{d}^{(J)})'\,\hat{\Lambda}^{-1}(\hat{P}^{(J)}c - \hat{d}^{(J)})$$

This estimator will be referred to as the two-step least squares estimator.

Discussion.

From the above considerations concerning the bias and variance of the elements of $\hat{\Lambda}$ it is clear that $\hat{\hat{\theta}}^{(J)}$ is not the minimum variance estimator $\tilde{\theta}^{(J)}$. However, numerical examples in the next section illustrate the close approximation of minimum variance in practice. The improvement over the one-step procedure is due to the appropriate choice of the elements of the weighting matrix Ω^{-1}. On the other hand the two-step procedure increases the computational effort since a larger number of finite Fourier transforms has to be computed, the least squares procedure needs to be carried out twice and in addition the covariance matrix of the residuals has to be estimated. This increase in computation time is not dramatic. Moreover, as the number of Fourier transforms to be computed for the two-step procedure increases, the use of FFT becomes progressively advantageous.

A careful reconsideration reveals that the above procedure for estimating the elements of Λ is partly based on intuitive arguments. In fact the tacit assumption has been made that half the sum of the squares of the measured residuals for a particular frequency is asymptotically an unbiased estimator of the asymptotic variance of the corresponding true residuals. In Section 3.6 it is shown that this is only true for those residuals which do not correspond to frequencies of non-zero harmonics of the test signal. Although the residuals corresponding to non-zero harmonics form a minority among the total number of residuals used for the computation of $\hat{\lambda}_w(\omega_{k_i})$ they give rise to a certain bias in $\hat{\lambda}_w(\omega_{k_i})$. Therefore in Section 3.6 a slightly modified estimator $\hat{\lambda}_{w'}(\omega_{k_i})$ is proposed defined by

$$\hat{\lambda}_{w'}(\omega_{k_i}) = \frac{JT}{2(2k_{w'}-k_h)} \sum_{\substack{k=k_iJ-k_{w'} \\ k/J \neq \text{ harmonic numbers} \\ \text{of non-zero harmonics}}}^{k_iJ+k_{w'}} \mathrm{Re}^2\{A(j2\pi k/JT)Z(k) +$$

$$-B(j2\pi k/JT)\exp(-j2\pi k\tau/JT)V(k)\}_{\theta = \hat{\theta}_I^{(J)}} +$$

$$+\mathrm{Im}^2\{A(j2\pi k/JT)Z(k) +$$

$$-B(j2\pi k/JT)\exp(-j2\pi k\tau/JT)V(k)\}_{\theta = \hat{\theta}_I^{(J)}}$$

where k_h denotes the number of non-zero harmonics of the test signal in the frequency interval $[(k_iJ - k_{w'})\ 2\pi/JT,\ (k_iJ + k_{w'})2\pi/JT]$. Note that the only difference between $\hat{\lambda}_w(\omega_{k_i})$ as defined by (1.3.9) and $\hat{\lambda}_{w'}(\omega_{k_i})$ is that in the expression for the latter the terms corresponding to non-zero harmonics of the test signal are not present. Two of the numerical examples in the next section are concerned with the two-step least squares estimator. In both cases the unmodified estimator $\hat{\lambda}_w(\omega_{k_i})$ has been used. The modified estimator $\hat{\lambda}_{w'}(\omega_{k_i})$ was developed later.

1.4. Numerical examples.

In this section the estimators discussed in the preceding sections are applied to computer generated data. The system used in all experiments is described by

$$0.25\,\frac{d^2y(t)}{dt^2} + 1.25\,\frac{dy(t)}{dt} + y(t) = u(t-\tau) \qquad (1.4.1)$$

The time constants of this system are 0.25 s and 1 s respectively. For comparison purposes the case of known and unknown time delay will be considered separately.

The test signals used are multifrequency binary signals (MFBS). These are periodic two-level signals which have the major part of their power concentrated in a relatively small number of widely spaced harmonics. Construction and properties of MFBS are discussed in Chapter 2. Two different MFBS are used in the experiments described below. The power of both signals is chosen equal to one. The first signal (MFBS1) has three dominant harmonics with harmonic numbers 1, 15 and 31. The power of these harmonics is $S_{uu}(\pm\, j2\pi/T) = 0.124$, $S_{uu}(\pm\, j2\pi\ 15/T) = 0.134$ and $S_{uu}(\pm\, j2\pi\ 31/T) = 0.118$ respectively. The sum equals 75% of the total power. The second signal (MFBS2) has five dominant harmonics with harmonic numbers 1, 7, 15, 23 and 31. The power of these harmonics is $S_{uu}(\pm\, j2\pi/T) = 0.059$, $S_{uu}(\pm\, j2\pi\ 7/T) = 0.090$, $S_{uu}(\pm\, j2\pi\ 15/T) = 0.076$, $S_{uu}(\pm\, j2\pi\ 23/T) = 0.084$ and $S_{uu}(\pm\, j2\pi\ 31/T) = 0.073$ respectively. The sum equals 76% of the total power.

MFBS are not to be confused with maximum length binary sequences (MLBS). As opposed to MFBS the power of MLBS is distributed over a large number of closely spaced harmonics. Moreover, since the envelope of the power density spectrum of MLBS is a $(\sin x/x)^2$ curve, it is more or less flat for low frequencies. For comparison purposes an MLBS is used in the numerical examples described below. It contains 63 steps in a period and its first 63 harmonics, representing 90% of its total power are taken into consideration. The total power is also chosen equal to one.

First experiment.

The system is described by (1.4.1). The time delay is assumed to be known and to be equal to zero. The input signal $u(t)$ is MFBS1. The frequency of $u(t)$ is 0.25 rad s^{-1}. The observations are

$$z(t) = y(t) + h(t)$$

The noise $h(t)$ is generated by

$$0.25 \frac{d^2h(t)}{dt^2} + 1.25 \frac{d}{dt} h(t) + h(t) = e(t)$$

where $e(t)$ is a zero-mean Gaussian process having power density spectrum

$$S_{ee}(j\omega) = \left(\frac{\sin 0.0245\ \omega}{0.0245\ \omega}\right)^2.$$

$S_{ee}(j\omega)$ is flat within 99% over the frequency range of the test signal. The signal-to-noise power ratio at the output with respect to the dominant harmonics is 0.61. Hundred independent records were generated consisting of four periods each. From these records the system coefficients were estimated for a record length of two and four periods respectively. The estimator was the weighted least squares estimator with weighting matrix $\Omega^{-1} = I$. The first column of Table 1 shows the average and the standard error of the estimates. For comparison purposes the second column shows the asymptotic standard deviation of the minimum variance least squares estimator in parentheses. The third column shows in parentheses the standard deviation of the minimum variance least squares estimator for the MLBS under the same conditions.

Table 1

$a_0 = 1.00$	$a_1 = 1.25$	$a_2 = 0.25$	
Record length: two periods			
$\hat{a}_0$	1.02 ± 0.32	(0.29)	(0.40)
$\hat{a}_1$	1.17 ± 0.32	(0.32)	(0.23)
$\hat{a}_2$	0.228 ± 0.065	(0.065)	(0.066)
Record length: four periods			
$\hat{a}_0$	1.00 ± 0.21	(0.20)	(0.29)
$\hat{a}_1$	1.22 ± 0.24	(0.22)	(0.17)
$\hat{a}_2$	0.241 ± 0.041	(0.046)	(0.047)

Discussion of the results of the first experiment.

According to (1.2.9) the elements of Λ are given by

$$\lambda(s_{k_i}) = \tfrac{1}{2}[\,|A(s_{k_i})|^2\{S_{h_2h_2}(s_{k_i}) + S_{h_3h_3}(s_{k_i})\} + $$
$$+ |B(s_{k_i})|^2 S_{g_2g_2}(s_{k_i})].$$

In this experiment $S_{h_3h_3}(s_{k_i}) = S_{g_2g_2}(s_{k_i}) = 0$ and $S_{h_2h_2}(s_{k_i}) = S_{ee}(s_{k_i})/|A(s_{k_i})|^2$. Since $S_{ee}(s_{k_i}) \approx 1$ over the frequency range of the test signal it follows that $\lambda(s_{k_i}) \approx \frac{1}{2}$. Thus $\Lambda(:)$diag $(1, 1, 1, 1, 1, 1) = I$. Therefore in this particular case the weighted least squares estimator with weighting matrix $\Omega^{-1} = I$ and the minimum variance least squares estimator coincide. Table 1 shows that even for a record length of two or four periods the standard error agrees with the asymptotic standard deviation of the minimum variance estimator.

The estimates of Table 1 are biased to some extent. However, the bias is small as compared to the standard error.

Furthermore it is observed that the bias is of first order. First order bias is bias inversely proportional to the observation time. First order bias may easily be removed using a bias elimination scheme due to Quenouille. See Kendall and Stuart (1967), page 5. Unfortunately this scheme involves an increase of the variance to first order and is therefore not suitable for records consisting of a small number of periods. On the other hand the bias of estimates obtained from records consisting of a large number of periods is small in comparison with the standard deviation and may therefore be neglected. Therefore application of the Quenouille scheme is advisable only if unbiasedness is essential and if many periods are available.

Finally, it is observed that in the case considered here the asymptotic standard deviations of the minimum variance least squares estimators for the MLBS and the MFBS are of the same order of magnitude.

Second experiment.

The system is described by (1.4.1). The time delay is assumed to be known and to be equal to zero. The input signal u(t) is MFBS1. The frequency of u(t) is 0.125 rad s^{-1}. The observations are

$$z(t) = y(t) + h(t)$$

The noise h(t) is a zero-mean Gaussian process having power density spectrum

$$S_{hh}(j\omega) = 0.036 \left(\frac{\sin 0.0490\ \omega}{0.0490\ \omega}\right)^2 .$$

$S_{hh}(j\omega)$ is flat within 99% over the frequency range of the test signal. The signal-to-noise power ratio at the output with respect to the dominant harmonics is 0.90. Hundred independent records were generated, each consisting of four periods. From these records the system coefficients were estimated for a record length of two and four periods respectively. The estimator was the weighted least squares estimator with weighting matrix $\Omega^{-1} = I$. The first column of Table 2 shows the average and standard error of the estimates. The second and third column show in parentheses the asymptotic standard deviation of the minimum variance least squares estimator for MFBS1 and the MLBS respectively under the same conditions.

Table 2

$a_0 = 1.00$		$a_1 = 1.25$	$a_2 = 0.25$
Record length: two periods			
$\hat{a}_0$	0.997 ± 0.047	(0.038)	(0.050)
$\hat{a}_1$	1.226 ± 0.131	(0.095)	(0.061)
$\hat{a}_2$	0.244 ± 0.060	(0.047)	(0.055)
Record length: four periods			
$\hat{a}_0$	0.998 ± 0.035	(0.027)	(0.035)
$\hat{a}_1$	1.237 ± 0.091	(0.067)	(0.043)
$\hat{a}_2$	0.246 ± 0.043	(0.033)	(0.039)

Discussion of the results of the second experiment.

As compared with the standard deviation of the minimum variance estimators, the estimates of Table 2 are less accurate than those of Table 1. This may be explained as follows. Recall that Λ is the weighting matrix of the minimum variance estimator. In this experiment Λ(:)diag (1, 1, 6, 6, 32, 32). The estimates of Table 2, however, have been computed using the least squares estimator with weighting matrix $\Omega^{-1} = I$. Note that in spite of the differences between the elements of Ω and Λ the estimates of Table 2 are still reasonably accurate.

Third experiment.

In this experiment the system coefficients are computed using the two-step procedure. The system and the test signal are the same as those of the preceding experiment. The half-width of the spectral window k_w, used in the second step, is equal to 6. The noise h(t) is a zero-mean Gaussian process having power density spectrum

$$S_{hh}(j\omega) = 0.033\left(\frac{\sin 0.0490\ \omega}{0.0490\ \omega}\right)^2$$

$S_{hh}(j\omega)$ is flat within 99% over the frequency range of the test signal. Fourty independent records were generated consisting of two periods each. The first column of Table 3 shows the average and standard error of the coefficient estimates computed in the first step. The second column shows the corresponding quantities computed in the second step. The third column shows in parentheses the asymptotic standard deviation of the minimum variance least squares estimator.

Table 3

a_0 = 1.00		a_1 = 1.25		a_2 = 0.25
$\hat{a}_0$	1.00 $\pm$ 0.05	$\hat{\hat{a}}_0$	1.00 $\pm$ 0.04	(0.04)
$\hat{a}_1$	1.21 $\pm$ 0.12	$\hat{\hat{a}}_1$	1.23 $\pm$ 0.10	(0.09)
$\hat{a}_2$	0.25 $\pm$ 0.06	$\hat{\hat{a}}_2$	0.24 $\pm$ 0.04	(0.05)

Discussion of the results of the third experiment.

The results of Table 3 show that the two-step procedure reduces the standard error of the estimates. Moreover, it is seen that in the case considered the standard error of the two-step estimates is comparable to the asymptotic standard deviation of the minimum variance least squares estimator.

Fourth experiment.

In this experiment both the system coefficients and the time delay are estimated using the two-step procedure. The system is described by (1.4.1) with $\tau = 0.7854$ s. For comparison purposes the noise and test signal characteristics are the same as those of the third experiment. The half-width of the spectral window k_w, used in the second step, is equal to 6. Fourty independent records were generated consisting of two periods each. The parameters were estimated by computing the system coefficients for $\tau = 0.7854 \pm k\ 0.0100$, where $k = 1, \ldots, 20$, and selecting the optimal solution. The first column of Table 4 shows the average and standard error of the parameter estimates computed in the first step. The second column shows the corresponding quantities computed in the second step. Finally, the third column shows in parentheses the asymptotic standard deviation of the minimum variance least squares estimator.

Table 4

$a_0 = 1.00$ $a_1 = 1.25$ $a_2 = 0.25$ $\tau = 0.7854$

$\hat{a}_0$	1.00 ± 0.05	$\hat{\hat{a}}_0$	1.00 ± 0.04	(0.04)
$\hat{a}_1$	1.20 ± 0.10	$\hat{\hat{a}}_1$	1.20 ± 0.10	(0.09)
$\hat{a}_2$	0.20 ± 0.16	$\hat{\hat{a}}_2$	0.23 ± 0.12	(0.13)
τ_d	0.83 ± 0.11	$\hat{\hat{\tau}}_d$	0.81 ± 0.10	(0.10)

Discussion of the results of the fourth experiment.

Again it is concluded that the two-step procedure improves the accuracy of the estimates and that the standard error of the two-step estimates is comparable with the asymptotic standard deviation of the minimum variance least squares estimator.

Finally, comparing the results of Table 3 and Table 4 it is important to note that the simultaneous estimation of the time delay causes a substantial increase of the standard error of the estimate of a_2.

Fifth experiment.

This experiment is an example of estimation in closed loop. The system is described by

$$0.25 \frac{d^2y(t)}{dt^2} + 1.25 \frac{d\,y(t)}{dt} + y(t) = u(t - 0.982)$$

The system is controlled by a three-term controller

$$\frac{d^2\,u(t)}{dt^2} + 18.1 \frac{d\,u(t)}{dt} = 1.485 \frac{d^2f(t)}{dt^2} + 30.5 \frac{df(t)}{dt} +$$

$$+ 11.1\ f(t)$$

where f(t) and u(t) are the input and output of the controller respectively. The controller input is

$$f(t) = s(t) - z(t)$$

where s(t) is the set point signal and z(t) is the measured system output defined by

$$z(t) = y(t) + h(t)$$

The disturbance h(t) is generated by

$$\frac{d\,h(t)}{dt} + h(t) = e(t)$$

where e(t) is a zero-mean Gaussian process having power density spectrum

$$S_{ee}(j\omega) = 0.049 \left(\frac{\sin 0.0245\ \omega}{0.0245\ \omega}\right)^2$$

$S_{ee}(j\omega)$ is flat within 99% over the frequency range of the test signal. The set point signal s(t) is BMFS2. The frequency of s(t) is 0.25 rad s^{-1}.

Seventeen independent records were generated consisting of four periods each. The columns of Table 5 show the average and the standard error for a record length of two and a record length of four periods respectively. The estimates were computed using the weighted least squares estimator with weighting matrix $\Omega = I$.

Table 5

$a_0 = 1.0$ $a_1 = 1.25$ $a_2 = 0.25$ $\tau = 0.982$

	Record length: two periods	Record length: four periods
$\hat{a}_0$	1.00 ± 0.07	0.99 ± 0.06
$\hat{a}_1$	1.30 ± 0.14	1.27 ± 0.07
$\hat{a}_2$	0.23 ± 0.02	0.24 ± 0.01
$\hat{\tau}$	0.99 ± 0.02	0.99 ± 0.01

Discussion of the results of the fifth experiment.

The results of Table 5 indicate that for increasing sample size the estimates converge to the actual values of the corresponding parameters. However, in view of the small sample size no far-reaching conclusions may be drawn. The estimates do not exhibit serious systematic deviations from the true values of the parameters.

1.5. Concluding remarks.

In the numerical examples of Section 1.4 the asymptotic standard deviation of the minimum variance least squares estimator has been adopted as a measure of accuracy of the estimates. In order to investigate as to in how far this is a sensible measure, it has to be compared with the minimum variance bound on the covariance of the estimator. In Section 3.5 it is shown that for Gaussian noise this bound is easily computed and coincides with the asymptotic covariance matrix of the minimum variance least squares estimator.

Other important problems not considered in Chapter 1 are the influence of the spectrum of the test signal on the accuracy of the estimator and the construction of test signals with specified spectral properties. These problems are discussed in Chapter 2.

CHAPTER 2

Selection and synthesis of periodic test signals.

The covariance matrix of the weighted least squares estimator of the system parameters proposed in this research was discussed in the previous chapter. The elements of this covariance matrix are functions of the weighting factors, the system parameters, the spectrum of the noise and the spectrum of the test signal. So for known system and noise characteristics the covariance matrix can be manipulated by selection of the test signal. The question then arises which test signal spectrum gives the most accurate estimates. Here the measure of error in estimation is taken as the weighted sum of the variances of the minimum variance least squares estimator. The particular spectrum and the corresponding test signals which minimize this measure are optimal in the defined sense. The particular choice of covariance matrix is motivated by the fact that it forms a lower bound for the covariance matrix of the weighted least squares estimator. Moreover, it is shown in Chapter 3 that this lower bound coincides with the minimum variance bound if the noise in the observations at the output obeys the normal distribution. Section 2.1 describes the functional relationship between the measure of estimation error and the spectrum of the test signal. This section also discusses the numerical procedure for the minimization of the measure. Section 2.2 computes numerically a number of optimal spectra and compares the corresponding covariance matrices with those for a maximum length binary sequence as well as for a particular multifrequency binary signal for the same system and noise.

The second part of this chapter is devoted to a search procedure for approximate synthesis of discrete interval binary multifrequency signals having specified spectra. Binary multifrequency signals are periodic two-level signals that have the major part of their power concentrated in a limited number of relatively widely spaced harmonics. The numerical synthesis procedure is described in Section 2.3. In Section 2.4 examples of signals computed using this procedure are given.

2.1 Computation of optimal test signal spectra.

According to (1.2.12) the asymptotic covariance matrix $\Sigma_{\tilde{\theta}\tilde{\theta}}$ of the minimum variance least squares estimator $\tilde{\theta}^{(J)}$ is described by

$$\Sigma_{\tilde{\theta}\tilde{\theta}} = \frac{1}{JT} (\Phi' \Lambda^{-1} \Phi)^{-1} \tag{2.1.1}$$

where J is the number of periods taken into consideration, T is the period of the periodic test signal u(t), while Φ is defined by (1.2.6) and (1.2.7) and Λ is defined by (1.2.8) and (1.2.9).

Now consider the particular case that u(t) is known. Let the observations at the output be

$$z(t) = y(t) + h(t)$$

where y(t) is the response to u(t) and where h(t) is a **stationary** stochastic process having power spectrum $S_{hh}(j\omega)$. Define $S_{uu}(j\omega_k) = |\gamma_{ku}|^2$ where γ_{ku} is the complex Fourier coefficient of the kth harmonic of u(t). Furthermore denote by $H_S(j\omega)$ the transfer function of the system and let

$$H_S(j\omega) = \frac{B(j\omega)}{A(j\omega)} \exp(-j\omega\tau)$$

Then it is shown in Section (3.4) that (2.1.1) may be written

$$\Sigma_{\tilde{\theta}\tilde{\theta}} = X^{-1} \tag{2.1.2}$$

where

$$x_{pq} = JT \sum_{\substack{k=-\infty \\ k \neq 0}}^{\infty} \frac{S_{uu}(j\omega_k)}{S_{hh}(j\omega_k)} |H_S(j\omega_k)|^2 \mathrm{Re}\{H_p(j\omega_k) H_q^{*}(j\omega_k)\} \tag{2.1.3}$$

where

$$H_i(j\omega) = \frac{(j\omega)^{i-1}}{A(j\omega)} \qquad i = 1, \ldots, N+1 \qquad (2.1.4)$$

$$H_i(j\omega) = - \frac{(j\omega)^{i-N-2}}{B(j\omega)} \qquad i = N+2, \ldots, N+M+1 \qquad (2.1.5)$$

and

$$H_i(j\omega) = j\omega \qquad i = N+M+2 \qquad (2.1.6)$$

where $\omega_k = 2\pi k/T$

It follows from (2.1.3) - (2.1.6) that for known system and noise characteristics the elements of X are known linear functions of the $S_{uu}(j\omega_k)$, $k = \pm 1, \pm 2, \ldots$ So the elements of X and therefore those of $\Sigma_{\tilde{\theta}\tilde{\theta}}$ can be manipulated by selection of the $S_{uu}(j\omega_k)$. Consequently, if some function of the elements of $\Sigma_{\tilde{\theta}\tilde{\theta}}$ is taken as a measure of the error in estimation of the parameters, this measure can in principle be minimized with respect to the $S_{uu}(j\omega_k)$.

Several measures of the error in multivariate estimation are known from the literature. A wellknown measure is the trace of the covariance matrix of the estimator involved. The determinant of the covariance matrix is also used. For a discussion of these measures and several others see Fedorov and Pázman (1968). Here the measure is taken as the weighted sum of the diagonal elements of the covariance matrix. This choice is motivated by the fact that it enables one to weight the various diagonal elements according to a desired accuracy. Furthermore, it is observed that in practice input power or output power is always restricted. The power constraint chosen in this research has the following form

$$\sum_{k=-\infty}^{\infty} \lambda_k \, S_{uu}(j\omega_k) = 1 \qquad (2.1.7)$$

where $\lambda_k \geqq 0$. Note that this is an input power constraint if $\lambda_k = 1$ for all k. Alternatively it is an output power constraint if $\lambda_k = |H_S(j\omega_k)|^2$

Summarizing, the problem considered in this section is to minimize the weighted sum of the diagonal elements of $\Sigma_{\tilde{\theta}\tilde{\theta}}$ with respect to $S_{uu}(j\omega_k)$, $k = \pm 1, \pm 2, \ldots$, subject to the constraint (2.1.7). The particular spectrum $\hat{S}_{uu}(j\omega_k)$ which corresponds to the minimum is defined as optimal. Test signals corresponding to the optimal spectrum will be referred to as optimal test signals.

It is observed that the weighted sum of the diagonal elements of $\Sigma_{\tilde{\theta}\tilde{\theta}}$ is nonlinear in the $S_{uu}(j\omega_k)$. As a result even in simple cases closed form expressions for the optimal spectrum are difficult to obtain. The computation of the gradient of the sum with respect to $S_{uu}(j\omega_k)$ is, however, relatively simple as is shown below. This offers the opportunity to apply powerful numerical optimization techniques.

Numerical computation of optimal spectra.

The weighted sum of the diagonal elements of $\Sigma_{\tilde{\theta}\tilde{\theta}}$ is described by

$$D = \sum_{i=1}^{N+M+2} \mu_i \sigma_{ii}$$

where σ_{ij} denotes the ij element of $\Sigma_{\tilde{\theta}\tilde{\theta}}$. Now the problem is to minimize D subject to the constraint (2.1.7) and the constraint

$$S_{uu}(j\omega_k) \geqq 0 \quad k = \pm 1, \pm 2, \ldots \tag{2.1.8}$$

First the constraints (2.1.7) and (2.1.8) are both removed by substituting

$$S_{uu}(j\omega_k) = \frac{y_k^2}{\sum\limits_{\substack{\ell=-\infty \\ \ell\neq 0}}^{\infty} \lambda_\ell y_\ell^2} \tag{2.1.9}$$

and minimizing D with respect to $y = (\ldots, y_{-k}, \ldots, y_{-1}, y_1, \ldots, y_k, \ldots)'$. The remainder of this section will be devoted to the derivation of a closed form expression for the gradient of D with respect to y. Once this expression has been obtained the application of numerical gradient techniques like the steepest descent method or the conjugate gradient method is straightforward.

First it is observed that in view of (2.1.9)

$$\frac{\partial D}{\partial y_k} = \sum_\ell \frac{\partial D}{\partial S_{uu}(j\omega_\ell)} \frac{\partial S_{uu}(j\omega_\ell)}{\partial y_k}$$

$$= \frac{2y_k}{\sum\limits_\ell \lambda_\ell y_\ell^2} \left(\frac{\partial D}{\partial S_{uu}(j\omega_k)} - \lambda_k \sum_\ell \frac{\partial D}{\partial S_{uu(j\omega_\ell)}} S_{uu}(j\omega_\ell) \right) \tag{2.1.10}$$

Furthermore, since $\Sigma_{\underset{\sim}{\theta}\underset{\sim}{\theta}} = X^{-1}$

$$D = \frac{1}{\det X} \sum_{i=1}^{N+M+2} \mu_i \det X_{ii} \tag{2.1.11}$$

where X_{pq} is the matrix obtained by eliminating the pth row and qth column from X. It follows from (2.1.11) that

$$\frac{\partial D}{\partial S_{uu}(j\omega_k)} = -\frac{1}{\det X}\frac{\partial \det X}{\partial S_{uu}(j\omega_k)} D +$$

$$+ \frac{1}{\det X}\sum_{i=1}^{N+M+2} \mu_i \frac{\partial \det X_{ii}}{\partial S_{uu}(j\omega_k)} \qquad (2.1.12)$$

The derivatives in this expression may be written

$$\frac{\partial \det X}{\partial S_{uu}(j\omega_k)} = \sum_{p=1}^{N+M+2}\sum_{q=1}^{N+M+2} (-1)^{p+q} \det X_{pq} \frac{\partial x_{pq}}{\partial S_{uu}(j\omega_k)} \qquad (2.1.13)$$

and

$$\frac{\partial \det X_{ii}}{\partial S_{uu}(j\omega_k)} = \sum_{m=1}^{M+N+1}\sum_{n=1}^{M+N+1} (-1)^{m+n} \det X_{ii,mn} \frac{\partial x_{ii,mn}}{\partial S_{uu}(j\omega_k)} \qquad (2.1.14)$$

where $X_{pq,mn}$ is the matrix obtained by eliminating the mth row and nth column from X_{pq} while $x_{pq,mn}$ is the mn element of X_{pq}. Note that the elements $x_{pq,mn}$ are a subset of the elements x_{pq}.

Now in order to obtain the gradient of D with respect to y at some point y^o the procedure is as follows. First the values of $S_{uu}(j\omega_k)$ are computed for y^o using (2.1.9). Then the elements of X are computed according to (2.1.3)-(2.1.6). Subsequent inversion of X using the Gauss elimination method yields detX and the cofactors $(-1)^{p+q}\det X_{pq}$. The derivatives $\frac{\partial x_{pq}}{\partial S_{uu}(j\omega_k)}$ also follow from (2.1.3)-(2.1.6). Furthermore since the x_{pq} are linear in the $S_{uu}(j\omega_k)$, the $\frac{\partial x_{pq}}{\partial S_{uu}(j\omega_k)}$ are independent of $S_{uu}(j\omega_k)$ and need to be computed only once. Now using (2.1.13) the first term of (2.1.12) is calculated. The second term of (2.1.12) is evaluated using (2.1.14). Inversion of the X_{ii} yields the cofactors $(-1)^{p+q}\det X_{ii,mn}$. The derivatives $\frac{\partial x_{ii,mn}}{\partial S_{uu}(j\omega_k)}$ are known since these are a subset of the $\frac{\partial x_{ii}}{\partial S_{uu}(j\omega_k)}$. Then using (2.1.14) the derivatives $\frac{\partial \det X_{ii}}{\partial S_{uu}(j\omega_k)}$, and hence the second term can be computed. This completes the computation of the gradient of D with respect to the $S_{uu}(j\omega_k)$. Finally, using (2.1.10) the gradient of D with respect to y is obtained.

2.2 Numerical examples.

The system considered is described by $a_2 \dfrac{d^2y(t)}{dt^2} + a_1 \dfrac{dy(t)}{dt} + a_o y(t) = u(t-\tau)$ where the values of the parameters are $a_2 = 0.25$, $a_1 = 1.25$, $a_o = 1$ and τ is arbitrary. The observations at the output are

$$z(t) = y(t) + h(t)$$

where $y(t)$ is the steady state reponse to the periodic test signal $u(t)$ and $h(t)$ is a stationary stochastic process having power spectral density

$$S_{hh}(j\omega) = S_o$$

where S_o is a constant. In what follows it is assumed that the observation time is an integral number of periods The power of all inputs is equal to one. Furthermore all weights in the measure of the estimation error are equal. So the measure is the trace of the covariance matrix $\Sigma_{\tilde{\theta}\tilde{\theta}}$.

Since $u(t)$ is periodic, time delay can only be estimated modulo the period of the fundamental. Therefore it is assumed that the time delay consists of the sum of a known integral multiple of the period of the fundamental and an unknown fraction of this period. Ambiguity in the interpretation of the estimates of this fraction may still arise if their standard deviation is comparable to the period of the fundamental. Therefore it is assumed that the standard deviation of the time delay estimates is small compared to the period of the fundamental Asymptotically this condition is always met.

In order to investigate the effect of estimation of time delay upon the variance of the estimator of the coefficients all computations were carried out both for

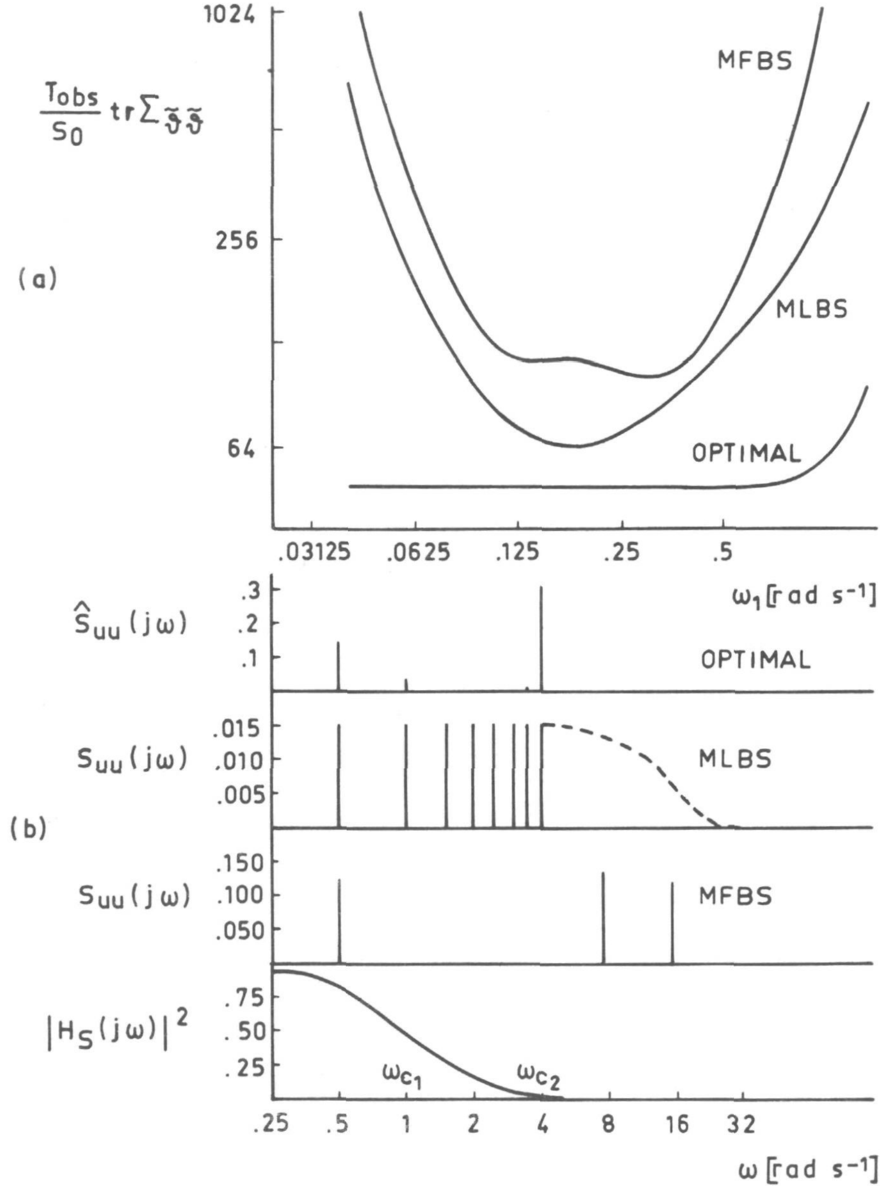

Figure 2.1.(a) The trace of the asymptotic covariance matrix of the minimum variance least squares estimator as a function of fundamental frequency for unknown time delay; (b) The optimal spectrum, the spectrum of the maximum length binary sequence, the spectrum of the multifrequency binary signal for a fundamental frequency of 0.5 rad s^{-1} and the square of the modulus of the system transfer function.

unknown and for known time delay. First the minimum values of the trace of $\Sigma_{\tilde{\theta}\tilde{\theta}}$ and the corresponding input spectra are determined for a number of fundamental frequencies of the input. For the same set of frequencies the trace of $\Sigma_{\hat{\theta}\hat{\theta}}$ is computed for a maximum length binary sequence (MLBS)and the multifrequency binary signal MFBS1 described in Section 1.4. The MLBS considered here has 63 steps in a period and only its first 63 harmonics, representing nearly 90% of its total power, are taken into consideration. The computations for the MFBS are based on its dominant harmonics only.

Figure 2.1 shows the results for joint estimation of coefficients and time delay. In figure 2.1(a) the trace of $\Sigma_{\tilde{\theta}\tilde{\theta}}$, normalized with respect to observation time T_{obs} and intensity of the noise, is plotted as a function of fundamental frequency. In what follows fundamental frequency will be denoted by ω_1.Figure 2.1(b) shows respectively the optimal spectrum, the spectrum of the MLBS, the spectrum of the MFBS for $\omega_1 = 0.5$ rad s^{-1} and the square of the modulus of the system transfer function. The corner frequencies of the system are indicated by ω_{c_1} and ω_{c_2}. Note that the bandwidth of the system and that of the optimal spectrum are approximately equal. Also note that the optimal test signal has four harmonics only. These are the fundamental, the second, the seventh and the eighth harmonic. With respect to the MLBS and the MFBS it follows from Figure 2.1(a) that these signals give results comparable to the optimum only for fundamental frequencies of about 0.2 rad s^{-1}. As an illustration the standard deviations of the minimum variance least squares estimator with the optimal input, the MLBS and the MFBS respectively are given in

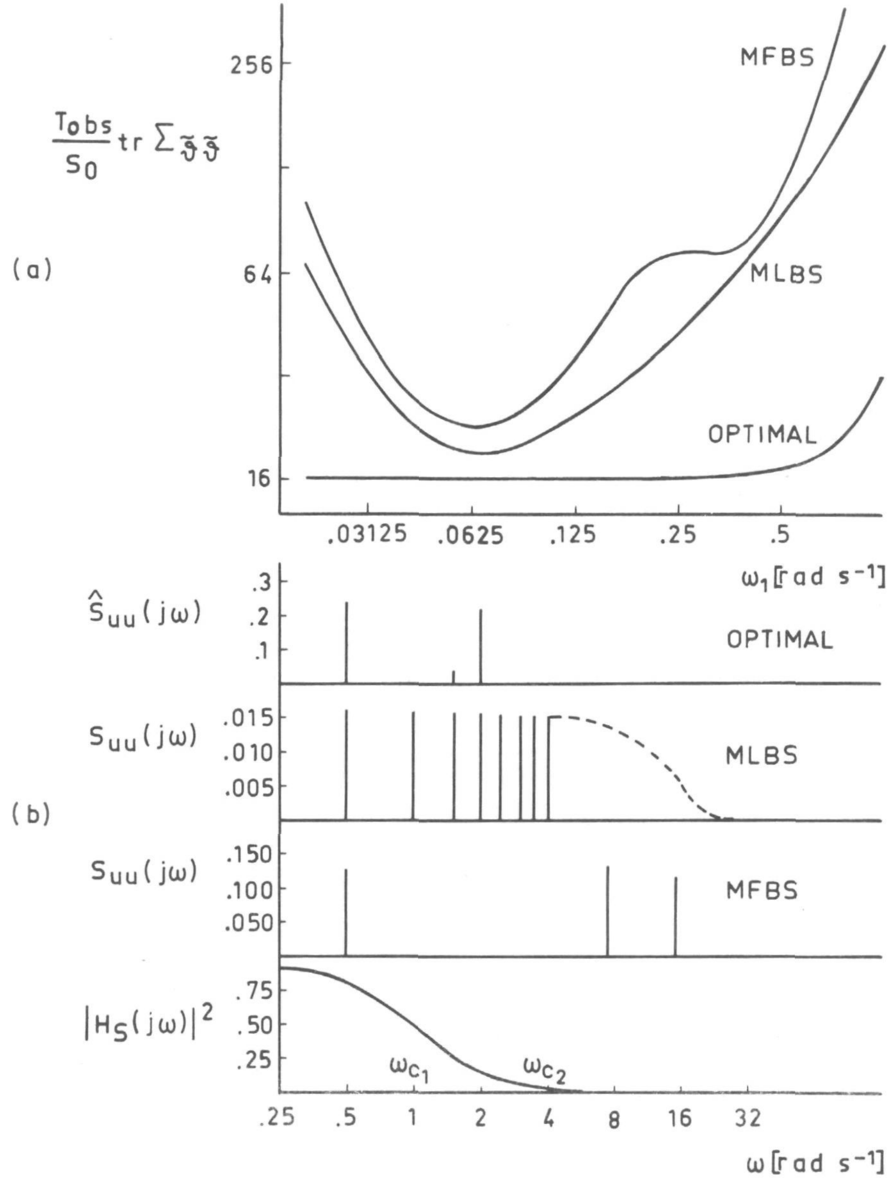

Figure 2.2.(a) The trace of the asymptotic covariance matrix of the minimum variance least squares estimator as a function of fundamental frequency for known time delay; (b) The optimal spectrum, the spectrum of the maximum-length binary sequence, the spectrum of the multifrequency binary signal for a fundamental frequency of 0.5 rad s^{-1} and the square of the modulus of the system transfer function.

Table 2.1 for $\omega_1 = 0.125$ rad s^{-1} and for $\omega_1 = 0.5$ rad s^{-1}. Note that for $\omega_1 = 0.5$ rad s^{-1} the deviations with the MLBS and MFBS from the optimum are quite serious.

Table 2.1

$\frac{T_{obs}}{S_o} \sigma$

$\omega_1 = 0.125$ rad s^{-1}

PARAMETER	OPTIMAL	MLBS	MFBS
a_o	2.50	2.75	2.04
a_1	3.88	3.71	5.12
a_2	4.50	5.71	7.44
τ	2.93	4.14	5.60

$\omega_1 = 0.5$ rad s^{-1}

PARAMETER	OPTIMAL	MLBS	MFBS
a_o	2.43	6.45	3.35
a_1	3.96	6.25	6.20
a_2	4.51	5.64	10.08
τ	2.93	3.44	4.17

Figure 2.2 and Table 2.2 representing the case of known time delay show corresponding results for joint estimation of the coefficients. Once more it is observed that MLBS and MFBS yield results comparable to the optimum only in a limited frequency range. Again the optimal spectrum, the spectrum of the MLBS, the spectrum of the MFBS and the square of the modulus of the system transfer function are shown for $\omega_1 = 0.5$ rad s^{-1}. Note that the bandwidth of the optimal spectrum is less than the system bandwidth.

Table 2.2

$$\frac{T_{obs}}{S_o}\,\sigma$$

$\omega_1 = 0.0625$ rad s^{-1}

PARAMETER	OPTIMAL	MLBS	MFBS
a_o	1.89	2.04	1.92
a_1	2.77	2.84	3.40
a_2	2.26	2.71	2.82

$\omega_1 = 0.5$ rad s^{-1}

PARAMETER	OPTIMAL	MLBS	MFBS
a_o	2.06	6.45	3.35
a_1	2.75	6.25	5.08
a_2	2.43	3.64	8.83

Comparing the results of Table 2.1 with those of Table 2.2 it is clear that the estimation of time delay causes an increase of the variance of some of the estimators of the remaining parameters. In particular this applies to the estimator of a_2. This is equivalent to the observation that the estimator of a_2 and that of τ are strongly covariant. For example, for $\omega_1 = 0.125$ rad s^{-1} the correlation coefficient of these estimators is as high as -0.81, -0.94 and -0.91 for respectively the optimal signal, the MFBS and the MLBS. This means that overestimation of τ usually yields underestimation of a_2 and vice versa. Although this may not be surprising it may easily give rise to an incorrect interpretation of the measurement results.

Concluding remarks

A practical conclusion to be drawn from the numerical examples is that in order to find an appropriate fundamental frequency it seems worthwile to carry out a priori computations of the minimum variance bound for a number of different fundamental frequencies of the input. These computations may be based on a priori knowledge about the system and noise obtained from mathematico-physical analysis or from previous experiments. In addition to information about the appropriate fundamental frequency such computations also provide information about possibly strong covariances between the estimators of the parameters.

2.3 Approximate synthesis of periodic binary signals having specified power spectra.

This section is concerned with the synthesis of discrete interval, binary multifrequency signals. These are discrete interval, two-level signals which have the major part of their power concentrated in a limited number of relatively widely spaced, selected harmonics. Since the signals are binary and discrete time they can be generated using simple digital circuitry and can easily be synchronized. A further advantage of these signals is that their peak factor, computed relative to the rms value of the selected harmonics, is relatively low as compared with that of signals obtained by adding a number of appropriate sinusoidal signals. The peak factor is defined as the ratio of the peak value to the rms value. The peak factor of a test signal should preferably be small since the range of linearity of the input transducer is usually limited.

Jensen (1959) constructed binary multifrequency signals by clipping the sum of a number of selected harmonics of appropriate amplitudes. It can be shown that this procedure maximizes the weighted sum of the Fourier coefficients of the resulting clipped signal, where the weighting factors are the corresponding Fourier coefficients of the original signal. See Van den Bos (1967). A disadvantage of this clipping procedure is that is not clear how the phases of the harmonics of the original signal should be chosen in order to obtain a clipped signal having the minimal peak factor relative to the selected harmonics. This difficulty is avoided by constructing binary signals which optimize a function of the power of the selected harmonics instead of a function of their Fourier coefficients. The remainder of this section discusses such a procedure.

In what follows the amplitude values of the discrete interval periodic binary signal u(t) are +1 and -1. So the power is equal to one. The number of steps in a period is I_s. The period is denoted by T. Zero crossings only occur at $t = iT/I_s$ where $i = 0, \ldots, I_s-1$. The value of u(t) on $iT/I_s < t < (i+1)T/I_s$ is denoted by u(i). Now let $S(j\omega_{\pm k_1}), \ldots, S(j\omega_{\pm k_L})$ denote the specified values of the power of the harmonics $\pm k_1, \ldots, \pm k_L$ respectively and let $\sum_{\substack{\ell=-L \\ \ell\neq 0}}^{L} S(j\omega_{k_\ell}) = 1$. The criterion to be minimized with respect to u(i), $i=0, \ldots, I_s-1$ is taken as

$$\sum_{\substack{\ell=-L \\ \ell\neq 0}}^{L} \{S(j\omega_{k_\ell}) - |\gamma_{k_\ell}|^2\}^2 \qquad (2.3.1)$$

Since the u(i) can only assume the values +1 and -1 and since (2.3.1) is nonlinear in the u(i), the minimization of (2.3.1) is a zero one integer nonlinear programming problem.

Is is clear that minimization of (2.3.1) via explicit enumeration of all possible combinations of values of the I_s variables u(i) is unattractive since for the signals considered here I_s is 64 or more. More efficient general procedures for the solution of zero one integer nonlinear programming are described in the literature. Plane and McMillan (1971) first reformulate the unconstrained nonlinear problem into a linear one by introducing a number of constraints and new variables and then compute the solution by a standard implicit enumeration algorithm. See Plane and McMillan (1971), Chapters 3 and 5. The main disadvantage of the method is that the size of the problem grows considerably. Using this method the problem at hand is transformed from an unconstrained nonlinear problem having I_s variables into a linear one having a number of constraints and a number of variables of the order I_s^4. Section 10.3 of Garfinkel and Nemhauser (1972) describes a method for solution of unconstrained zero one integer nonlinear programming problems. It can be shown that for the particular problem considered this method hardly reduces the computational effort as compared with explicit enumeration of all possible combinations of values.

Since no more suitable algorithm for minimization of (2.3.1) was known to the author, a computationally simple, heuristic procedure was developed. This procedure is extensively described in Van den Bos (1967) and may be summarized as follows. First an initial configuration is generated by assigning at random the value +1 or -1 to each u(i). This is done for each u(i) independently. Next the Fourier coefficients and the corresponding criterion are computed. Now, using a uniform random number generator a number between 0 and I_s-1 is produced. Let this number be i_o. Then $u(i_o)$ is converted into $-u(i_o)$ and next the Fourier coefficients and the corresponding criterion are computed for this new configuration. If the inversion of $u(i_o)$ is an improvement in the sense of the criterion it is maintained, if not the original situation is restored. Next the same computations are carried out for $u(i_o+1)$ and so on up to $u(i_o+I_s-1)$ where the argument should be taken modulo I_s. Now using the uniform random number generator a new starting point i_1 is generated and a second run is made. This process is continued until a complete run without further improvement occurs. This procedure does not necessarily yield the optimal solution, but by repeating it the probability of finding the optimal solution or a suffiently good suboptimal solution, greatly increases. Therefore the procedure is repeated a number of times and the best solution is selected.

A number of signals computed using this procedure is given in the next section.

2.4 Numerical examples

In all numerical examples described in this section the specified values of the power of the L selected harmonics are $S(j\omega_{k_\ell}) = 0.5/L$ where $\ell = \pm 1, \ldots, \pm L$. The notation of the signals is illustrated by the following examples. A signal that is +1 on two consecutive intervals and -1 on the next five is denoted 2^+5^-. The efficiency of a signal is defined as the ratio of the power concentrated in the selected harmonics to the total power.

Signal 1

$I_s = 512$. L=6 . $k_1=1$, $k_2=2$, $k_3=4$, $k_4=8$, $k_5=16$ and $k_6=32$.
The signal is symmetric about t=0.
The efficiency is 0.702. For $0 < t < T/2$ the signal is described by $23^+5^-9^+3^-45^+25^-5^+27^-3^+27^-11^+3^-27^+24^-7^+12^-$

Selected harmonics	α_{k_ℓ}	β_{k_ℓ}
1	+0.241	-
2	+0.239	-
4	-0.251	-
8	+0.239	-
16	-0.241	-
32	+0.247	-

Signal 2

$I_s = 64$. L=2 . $k_1=1$ and $k_2=6$. The signal is symmetric about t=0. The efficiency is 0.645. For $0 < t < T/2$ the signal is described by $9^-4^+5^-8^+2^-1 4^+$.

Selected harmonics	α_{k_ℓ}	β_{k_ℓ}
1	-0.398	-
6	+0.406	-

Signal 3

$I_s = 512$. $L = 3$. $k_1 = 1$, $k_2 = 15$ and $k_3 = 31$. The signal is skew symmetric about $t = 0$ and symmetric about $T/4$. The efficiency is 0.751. For $0 < t < T/4$ the signal is described by $11^- 4^+ 8^- 10^+ 25^- 7^+ 26^- 7^+ 27^- 3^+$.

Selected harmonics	α_{k_ℓ}	β_{k_ℓ}
1	-	-0.352
15	-	-0.366
31	-	-0.343

Signal 4

$I_s = 256$. $L = 5$. $k_1 = 1$, $k_2 = 7$, $k_3 = 15$, $k_4 = 23$ and $k_5 = 31$.
The signal is skew symmetric about $t = 0$ and symmetric about $T/4$. The efficiency is 0.760. For $0 < t < T/4$ the signal is described by $5^- 12^+ 4^- 1^+ 6^- 6^+ 4^- 19^+ 4^- 3^+$.

Selected harmonics	α_{k_ℓ}	β_{k_ℓ}
1	-	+0.243
7	-	+0.299
15	-	-0.276
23	-	-0.289
31	-	-0.270

Signal 5

$I_s = 256$. $L = 5$. $k_1 = 2$, $k_2 = 8$, $k_3 = 16$, $k_4 = 24$ and $k_5 = 32$.
The signal is symmetric about $t = 0$ and about $T/4$. The efficiency is 0.697. For $0 < t < T/4$ the signal is described by $2^- 5^+ 3^- 11^+ 5^- 4^+ 5^- 2^+ 7^- 8^+ 12^-$.

Selected harmonics	α_{k_ℓ}	β_{k_ℓ}
2	0.246	-
8	0.294	-
16	0.263	-
24	0.269	-
32	-0.272	-

Concluding remarks

In the above numerical examples only signals having special symmetry properties have been computed. These symmetry properties result in considerable savings in computation time. For example, for the signals 1 and 2 the search procedure needs to be carried out for the first half of the period only. In the case of the signals 3, 4 and 5 only the first quarter of the period needs to be taken into consideration. On the other hand the introduction of symmetry properties restricts the class of possible phases of the selected harmonics considerably. This implies that possibly a number of signals which are better in the sense of the criterion is precluded.

Since the signals are binary, their peak factor relative to the rms value of the selected harmonics is equal to the square root of the reciprocal of the efficiency. For example, the peak factor of signal 2 is 1.24. Now for comparison purposes consider the signal $\cos(2\pi t/T) + \cos(2\pi 6/T + \varphi)$. The peak factor of this signal can be minimized by adjusting the phase angle φ. The minimum value is 1.97. So in spite of the fact that signal 2 has only 64.5% of its power concentrated in the selected harmonics, its peak factor, computed relative to the rms factor of the selected harmonics, is relatively small.

Finally observe that signal 4 contains only odd harmonics. Signal 5, however, is composed of even harmonics only. So these signals have no common harmonics and are therefore suitable for application to systems having more than one input.

CHAPTER 3

Proofs and mathematical details

This chapter contains proofs and mathematical details of some results discussed in the previous chapters.

The estimator of the system parameters proposed in Chapter 1 is a function of the estimator of the Fourier coefficients of input and output. The estimator of the Fourier coefficients is taken as the least squares estimator. This choice is motivated by the fact that this estimator requires little a priori knowledge about the noise. Detailed a priori knowledge about the noise required for more sophisticated estimators is usually not available. Moreover, the least squares estimator is computationally convenient. Section 3.1 discusses this estimator and establishes sufficient conditions for its consistency.

The results of Section 3.1 are used in Section 3.2 where the expression for the least squares estimator of the system parameters is derived. This section also discusses the conditions for the consistency of the estimator of the system parameters.

Since the estimator of the system parameters is a differentiable function of the estimator of the Fourier coefficients, its asymptotic covariance matrix can be computed from the asymptotic covariance matrix of the Fourier coefficients using a theorem due to Goldberger. For that purpose first the asymptotic covariance matrix of the estimator of the Fourier coefficients is computed in Section 3.3. Then using the results of Section 3.3, Section 3.4 derives the expression for the asymptotic covariance of the estimator of the system parameters. Section 3.4 also computes the asymptotic covariance matrix of the residuals.

Section 3.5 computes the minimum variance bound for the estimation of the system parameters. This section also shows that the minimum variance least squares estimator discussed in Section 1.2 achieves the minimum variance bound asymptotically if the noise in the observations at the output is normally distributed.

Finally, in Section 3.6 the asymptotic expectation of the estimator of the weighting matrix used in the two-step least squares procedure is computed. It turns out that this estimator, which was selected on intuitive grounds, is slightly biased. It is shown that by a minor modification of the estimator the bias is easily removed.

3.1. Least squares estimation of Fourier coefficients

This section discusses the least squares estimation of the Fourier coefficients of a periodic function disturbed by additive noise. Sufficient conditions for the consistency of this estimator are given.

Let f(t) be periodic with period T. Define the complex Fourier coefficient of the kth harmonic of f(t) by

$$\gamma_{kf} = \frac{1}{T} \int_0^T f(t)\exp(-j2\pi kt/T)dt \tag{3.1.1}$$

and denote

$$\gamma_{kf} = \alpha_{kf} - j\beta_{kf}$$

Let

$$w(t) = f(t) + n(t)$$

be observed, where n(t) is a stationary stochastic process having an autocovariance function $R_{nn}(t_1)$. Furthermore let w(t) be observed for $0 \leq t < JT$ where J is an integer. The least squares estimator $\hat{\gamma}_{kf}^{(J)}$ of γ_{kf} is that value of γ_{kf} which minimizes

$$\int_0^{JT} \{w(t) - \sum_{i=-\infty}^{\infty} \gamma_{if} \exp(j2\pi it/T)\}^2 dt$$

Hence

$$\hat{\gamma}_{kf}^{(J)} = \frac{1}{JT} \int_0^{JT} w(t)\exp(-j2\pi kt/T)dt \tag{3.1.2}$$

or, equivalently

$$\hat{\alpha}_{kf}^{(J)} = \frac{1}{JT} \int_0^{JT} w(t)\cos 2\pi kt/T \; dt \tag{3.1.3}$$

and

$$\hat{\beta}_{kf} = \frac{1}{JT} \int_0^{JT} w(t)\sin 2\pi kt/T \; dt \tag{3.1.4}$$

The estimators (3.1.2)- (3.1.4) are unbiased since

$$E\hat{\gamma}_{kf}^{(J)} = E\,\frac{1}{JT}\int_0^{JT} w(t)\exp(-j2\pi kt/T)dt$$

$$= \frac{1}{JT}\int_0^{JT} Ew(t)\exp(-j2\pi kt/T)dt$$

$$= \frac{1}{JT}\int_0^{JT} (f(t) + \mu\ )\exp(-j2\pi kt/T)dt$$

$$= \frac{1}{JT}\int_0^{JT} f(t)\exp(-j2\pi kt/T)dt = \gamma_{kf}$$

where $\mu\ = En(t)$.

In order to investigate the consistency of $\hat{\gamma}_{kf}^{(J)}$ first consider the random variable $\sqrt{JT}\ \hat{\alpha}_{kf}^{(J)}$. The variance of this random variable is

$$E(\sqrt{JT}\ \hat{\alpha}_{kf}^{(J)} - E\sqrt{JT}\ \hat{\alpha}_{kf}^{(J)})^2 = E\ JT(\hat{\alpha}_{kf}^{(J)} - \alpha_{kf})^2$$

$$= \frac{1}{JT}\ E(\int_0^{JT} n(t)\cos\,2\pi kt/T\ dt)^2$$

$$= \frac{1}{JT}\int_0^{JT} dt_1 \int_0^{JT} dt_2\ R_{nn}(t_2 - t_1)\cos\,2\pi kt_1/T\,\cos\,2\pi kt_2/T$$

$$= \frac{1}{JT}\int_0^{JT} dt\ R_{nn}(t)\int_0^{JT-t} dt_1\ \cos\,2\pi kt_1/T\,\cos\,2\pi k(t_1 + t)/T +$$

$$+ \frac{1}{JT}\int_{-JT}^{0} dt\ R_{nn}(t)\int_{-t}^{JT} dt_1\ \cos\,2\pi kt_1/T\,\cos\,2\pi k(t_1 + t)/T$$

$$= \frac{1}{2}\int_{-JT}^{JT} R_{nn}(t)(1 - |t|/JT)\cos\,2\pi kt/T\ dt +$$

$$- \frac{1}{8\pi kJ}\int_{-JT}^{JT} R_{nn}(t)\ \sin\,2\pi k|t|/T\ dt \qquad (3.1.5)$$

Now assume that $R_{nn}(t)$ is absolutely integrable. This implies that

$$\int_{-\infty}^{\infty} |R_{nn}(t)|\ dt < \infty \qquad (3.1.6)$$

Then

$$\int_{-JT}^{JT} R_{nn}(t)(1 - |t|/JT)\cos 2\pi kt/T \, dt$$

$$\leqslant \int_{-JT}^{JT} |R_{nn}(t)(1 - |t|/JT)\cos 2\pi kt/T| dt$$

$$\leqslant \int_{-\infty}^{+\infty} |R_{nn}(t)| \, dt < \infty \tag{3.1.7}$$

and

$$\int_{-JT}^{JT} R_{nn}(t) \sin 2\pi k|t|/T \, dt$$

$$\leqslant \int_{-JT}^{JT} |R_{nn}(t) \sin 2\pi k|t|/T| dt$$

$$\leqslant \int_{-\infty}^{\infty} |R_{nn}(t)| dt < \infty \tag{3.1.8}$$

It then follows from (3.1.5), (3.1.7) and (3.1.8) that

$$E \; JT(\hat{\alpha}_{kf}^{(J)} - \alpha_{kf})^2 < \infty$$

and hence

$$\lim_{J\to\infty} E(\hat{\alpha}_{kf}^{(J)} - \alpha_{kf})^2 = 0$$

It is concluded that $\hat{\alpha}_{kf}^{(J)}$ converges in the mean square to α_{kf}. Consequently $\hat{\alpha}_{kf}^{(J)}$ converges in probability to α_{kf}. Therefore the condition that $R_{nn}(t)$ is absolutely integrable is a sufficient condition for the consistency of $\hat{\alpha}_{kf}^{(J)}$. If in addition it is assumed that

$$\int_{-\infty}^{\infty} |t \; R_{nn}(t)| \, dt < \infty \tag{3.1.9}$$

it follows from (3.1.5) that

$$\lim_{J\to\infty} E\left(\sqrt{JT}\,\hat{\alpha}_{kf}^{(J)} - E\,\sqrt{JT}\,\hat{\alpha}_{kf}^{(J)}\right)^2$$

$$= \frac{1}{2}\int_{-\infty}^{\infty} R_{nn}(t)\cos 2\pi kt/T\,dt = \frac{1}{2}S_{nn}(j2\pi k/T) \qquad (3.1.10)$$

where $S_{nn}(j\omega)$ denotes the power density spectrum of $n(t)$. In the same way it can be shown that (3.1.6) is a sufficient condition for the consistency of $\hat{\beta}_{kf}^{(J)}$ and that if both (3.1.6) and (3.1.9) are met

$$\lim_{J\to\infty} E\left(\sqrt{JT}\,\hat{\beta}_{kf}^{(J)} - E\,\sqrt{JT}\,\hat{\beta}_{kf}^{(J)}\right)^2 =$$

$$= \frac{1}{2}S_{nn}(j2\pi k/T) \qquad (3.1.11)$$

The condition (3.1.6) for the consistency of $\hat{\alpha}_{kf}^{(J)}$ and $\hat{\beta}_{kf}^{(J)}$ is a sufficient condition. The following example shows that it is not a necessary condition. For example, consider the case that $n(t)$ is a periodic process. A periodic process is a process having a periodic autocovariance function. See Van Trees (1968), page 209. An example of a periodic autocovariance function is

$$R_{nn}(t) = \cos(2\pi t/T') \qquad (3.1.12)$$

where T' denotes the period. Clearly this autocovariance function is not absolutely integrable. Substituting (3.1.12) in (3.1.5) yields

$$E\left(\sqrt{JT}\,\hat{\alpha}_{kf}^{(J)} - E\,\sqrt{JT}\,\hat{\alpha}_{kf}^{(J)}\right)^2$$

$$= \frac{1}{2}\int_{-JT}^{JT}\cos(2\pi t/T')(1 - |t|/JT)\cos 2\pi kt/T\,dt$$

$$- \frac{1}{8\pi kJ}\int_{-JT}^{JT}\cos(2\pi t/T')\sin 2\pi k|t|/T\,dt$$

$$= \frac{JT}{4}\,[\sin\{\pi(k/T + 1/T')JT\}/\{\pi(k/T + 1/T')JT\}]^2 +$$

$$+ \frac{JT}{4}\,[\sin\{\pi(k/T - 1/T')JT\}/\{\pi(k/T - 1/T')JT\}]^2$$

$$- \frac{1}{8\pi kJ} \int_0^{JT} \sin\,\{2\pi(k/T + 1/T')t\}dt$$

$$- \frac{1}{8\pi kJ} \int_0^{JT} \sin\,\{2\pi(k/T - 1/T')t\}dt$$

Hence

$$\lim_{J\to\infty} (\sqrt{JT}\,\hat{\alpha}_{kf}^{(J)} - E\,\sqrt{JT}\,\hat{\alpha}_{kf}^{(J)})^2 = \begin{cases} \infty \text{ for } 1/T' = k/T \\ 0 \text{ for } 1/T' \neq k/T \end{cases}$$

So in spite of the fact that the autocovariance function of the periodic process $n(t)$ is not absolutely integrable, $\hat{\alpha}_{kf}^{(J)}$ converges in the mean square to α_{kf} provided that the frequency $2\pi k/T$ does not coincide with the frequency $2\pi/T'$ of the process $n(t)$. Correspondingly,

$$\lim_{J\to\infty} (\sqrt{JT}\,\hat{\beta}_{kf}^{(J)} - E\,\sqrt{JT}\,\hat{\beta}_{kf}^{(J)})^2 = \begin{cases} \infty \text{ for } 1/T' = k/T \\ 0 \text{ for } 1/T' \neq k/T \end{cases}$$

3.2. The estimator of the system parameters and its consistency

In this section first the expression for the least squares estimator of the system parameters is derived. The conditions for its consistency are derived next.

According to Section 1.1 the least squares estimator $\hat{c}^{(J)}$ is that solution for c which minimizes

$$(\hat{P}^{(J)}c - \hat{d}^{(J)})' \, \Omega^{-1} \, (\hat{P}^{(J)}c - \hat{d}^{(J)}) \tag{3.2.1}$$

The gradient of (3.2.1) with respect to c is

$$\frac{\partial}{\partial c} (\hat{P}^{(J)}c - \hat{d}^{(J)})' \, \Omega^{-1} \, (\hat{P}^{(J)}c - \hat{d}^{(J)})$$

$$= 2\hat{P}'^{(J)}\Omega^{-1} \, \hat{P}^{(J)}c - 2\hat{P}'^{(J)}\Omega^{-1} \, \hat{d}^{(J)} \tag{3.2.2}$$

So $\hat{c}^{(J)}$ must satisfy

$$\hat{P}'^{(J)}\Omega^{-1} \, \hat{P}^{(J)}\hat{c}^{(J)} - \hat{P}'^{(J)}\Omega^{-1} \, \hat{d}^{(J)} = 0 \tag{3.2.3}$$

and hence

$$\hat{c}^{(J)} = (\hat{P}'^{(J)}\Omega^{-1} \, \hat{P}^{(J)})^{-1} \, \hat{P}'^{(J)}\Omega^{-1} \, \hat{d}^{(J)} \tag{3.2.4}$$

where it is assumed that $\hat{P}'^{(J)}\Omega^{-1} \, \hat{P}^{(J)}$ is nonsingular. In order to investigate the validity of this assumption first the rank of the matrix P is determined.

The rank of P

The rank of a matrix is defined as the number of vectors in the largest linearly independent set of vectors which can be constructed from the columns of that matrix. The matrix P is $2L \times (N+M+1)$. Now assume that $2L \geqslant N+M+1$. This implies that the number of system coefficients is supposed to be smaller than or equal to twice the number of harmonics taken into consideration. So $r(P) \leqslant N+M+1$, where $r(P)$ denotes the rank of P. The matrix P is defined by (1.1.6) and (1.1.7).

From these equations it follows that the columns of P are linearly dependent if there exists a set of scalars μ_o, ..., μ_N, ν_o, ..., ν_{M-1}, not all of which are zero, such that

$$\sum_{n=o}^{N} \mu_n \operatorname{Re} s_{k_\ell}^n \gamma_{k_\ell y} - \sum_{m=0}^{M-1} \nu_m \operatorname{Re} s_{k_\ell}^m \gamma_{k_\ell u} = 0 \qquad (3.2.5)$$

and

$$\sum_{n=0}^{N} \mu_n \operatorname{Im} s_{k_\ell}^n \gamma_{k_\ell y} - \sum_{m=0}^{M-1} \nu_m \operatorname{Im} s_{k_\ell}^m \gamma_{k_\ell u} = 0 \qquad (3.2.6)$$

for $\ell = 1, \ldots, L$. The real equations (3.2.5) and (3.2.6) are equivalent to the complex equation

$$D(s_{k_\ell})\gamma_{k_\ell y} - E(s_{k_\ell})\gamma_{k_\ell u} = 0 \qquad (3.2.7)$$

where

$$D(s) = \mu_o + \mu_1 s + \ldots + \mu_N s^N$$

and

$$E(s) = \nu_o + \nu_1 s + \ldots + \nu_{M-1} s^{M-1}$$

for $\ell = 1, \ldots, L$. Substituting $\gamma_{k_\ell u} B(s_{k_\ell})/A(s_{k_\ell})$ for $\gamma_{k_\ell y}$ in (3.2.7) yields

$$\{D(s_{k_\ell}) B(s_{k_\ell}) - E(s_{k_\ell}) A(s_{k_\ell})\}\gamma_{k_\ell u} = 0 \qquad (3.2.8)$$

since $A(s_{k_\ell}) \neq 0$ for $\ell = 1, \ldots, L$. Now let $\gamma_{k_\ell u} \neq 0$ for $\ell = 1, \ldots, L$. Then it follows from (3.2.8) that the columns of P are linearly dependent if a set of scalars μ_o, ..., μ_N, ν_o, ..., ν_{M-1} can be found such that

$$\operatorname{Re}\{D(s_{k_\ell}) B(s_{k_\ell}) - E(s_{k_\ell}) A(s_{k_\ell})\} = 0 \qquad (3.2.9)$$

and

$$\mathrm{Im}\{D(s_{k_\ell})\, B(s_{k_\ell}) - E(s_{k_\ell})\, A(s_{k_\ell})\} = 0 \qquad (3.2.10)$$

for $\ell = 1, \ldots, L$. Now suppose that such a set of scalars exists. Taking into account that $s_{k_\ell} = j\omega_{k_\ell}$ it then follows from (3.2.9) and (3.2.10) that

$$\omega = \omega_{k_1}, \omega_{k_2}, \ldots, \omega_{k_L} \qquad (3.2.11)$$

satisfy

$$\mathrm{Re}\{D(j\omega)\, B(j\omega) - E(j\omega)\, A(j\omega)\} = 0 \qquad (3.2.12)$$

and

$$\mathrm{Im}\{D(j\omega)\, B(j\omega) - E(j\omega)\, A(j\omega)\} = 0 \qquad (3.2.13)$$

The left-hand members of (3.2.12) and (3.2.13) are even and odd polynomials in ω respectively. Consequently

$$\omega = -\omega_{k_1}, -\omega_{k_2}, \ldots, -\omega_{k_L}$$

must also satisfy (3.2.12) and (3.2.13). So both (3.2.12) and (3.2.13) have at least 2L roots. According to the definition of the polynomials $A(j\omega)$, $B(j\omega)$, $D(j\omega)$ and $E(j\omega)$ neither the degree of (3.2.12) nor that of (3.2.13) can exceed N+M. Hence neither the number of roots of (3.2.12) nor that of (3.2.13) can exceed N+M. So if the set of scalars exists the inequality $N+M \geqslant 2L$ holds. Since by assumption, however, $2L \geqslant N+M+1 > N+M$ it is concluded that the required set of scalars does not exist. Hence the columns of P are linearly independent and $r(P) = N+M+1$ if $N+M+1 \leqslant 2L$.

Returning to (3.2.4) consider the matrix $P'\Omega^{-1}P$. Let $2L \geqslant N+M+1$ and let $\gamma_{k_\ell u} \neq 0$ for $\ell=1, \ldots, L$. Then P is a $2L \times (N+M+1)$ matrix having rank $r(P) = N+M+1$. By definition Ω^{-1} is a $2L \times 2L$ positive definite matrix. Then $P'\Omega^{-1}P$ is positive definite and therefore nonsingular. See Goldberger

(1964), page 35. Consequently $(P'\Omega^{-1}P)^{-1}$ exists and therefore generally $(\hat{P}'^{(J)}\Omega^{-1}\,\hat{P}^{(J)})^{-1}$ exists. Then the elements of

$$\hat{c}^{(J)} = (\hat{P}'^{(J)}\Omega^{-1}\,\hat{P}^{(J)})^{-1}\,\hat{P}'^{(J)}\Omega^{-1}\,\hat{d}^{(J)}$$

are continuous functions of the estimators of the Fourier coefficients $\hat{\alpha}^{(J)}_{k_\ell u}$, $\hat{\beta}^{(J)}_{k_\ell u}$, $\hat{\alpha}^{(J)}_{k_\ell y}$ and $\hat{\beta}^{(J)}_{k_\ell y}$. A theorem due to Slutsky states that the probability limit of a continuous function is the function of the probability limits of the arguments. See Wilks (1962), page 102. Now assume that the estimators of the Fourier coefficients are consistent. If then follows from Slutsky's theorem that

$$\text{p lim } \hat{c}^{(J)} = \text{p lim } (\hat{P}'^{(J)}\Omega^{-1}\,\hat{P}^{(J)})^{-1}\,\hat{P}'^{(J)}\Omega^{-1}\,\hat{d}^{(J)}$$

$$= (P'\Omega^{-1}P)^{-1}\,P'\Omega^{-1}d \qquad (3.2.14)$$

where p lim $\hat{c}^{(J)}$ denotes the probability limit of the sequence $\{\hat{c}^{(J)}\}$. According to (1.1.10)

$$Pc = d \qquad (3.2.15)$$

Finally, substitution of (3.2.15) in (3.2.14) yields

$$\text{p lim } \hat{c}^{(J)} = (P'\Omega^{-1}P)^{-1}\,P'\Omega^{-1}Pc = c$$

From the above considerations it is concluded that $\hat{c}^{(J)}$ is a consistent estimator of c if the following conditions are both met:

1) the estimator of the Fourier coefficients is consistent
2) the number of unknown system coefficients is smaller than or equal to twice the number of non-zero harmonics taken into consideration.

Consistency in the case of unknown time delay

As described in Section 1.1 the estimation procedure for the case of unknown time delay is as follows. The coefficients of the differential equation are estimated for a number of selected values of time delay. The optimal value of the time delay and the corresponding solution for the coefficients are selected.

It is observed that this estimation procedure can only be consistent if the selected range of values of time delay includes the true value. Furthermore as the test signal is periodic with period T the time delay can only be computed modulo T. Therefore in this research it is assumed that the time delay consists of the sum of a known integral multiple of T and an unknown fraction of T. For simplicity in what follows this unknown fraction will be referred to as time delay. Note that with this definition the value of the time delay always lies within the interval $[0, T)$. Consequently, is a sufficiently large number of equidistant values of the time delay is taken into consideration, these values always include the true value or at least a value close to the true value. In what follows it is supposed that this condition is met.

First consider the case that the Fourier coefficients of the test signal and those of the response are exactly known. Now select from (3.2.15) N+M+2 equations. These equations constitute N+M+2 relations between N+M+2 unknowns. So in general the coefficients and the delay can uniquely be determined if $N+M+2 \leqslant 2L$, 2L being the number of equations in (3.2.15). Note that for unambiguous determination of the time delay the sum of the harmonics to which the N+M+2 selected equations correspond may not repeat itself within one period of the fundamental. This difficulty is avoided by including one or both equations corresponding to the fundamental frequency.

Next consider the case that only estimates of the Fourier coefficients of the test signal and the response are available. First it is observed that the estimators of the system coefficients and the time delay used in this research are continuous functions of the estimators of the Fourier coefficients. Now recall that Slutsky's theorem states that the probability limit of a continuous function is the function of the probability limits. So, if the estimators of the Fourier coefficients are consistent, the estimator of the system parameters converges in probability to the solution for the system parameters in the noiseless case. It has been shown above that this solution is in general the true value of the parameters if $2L \geqslant N+M+2$. Note, however, that ambiguity in the interpretation of time delay estimates may arise if the standard deviation of the time delay estimates is comparable to T. Therefore it is assumed that the standard deviation is small compared with T. Asymptotically this condition is always met.

It is concluded that the proposed least squares estimator for the system coefficients and the time delay is in general consistent if the following conditions are both satisfied:

1) the estimator of the Fourier coefficients is consistent
2) the number of unknown system parameters is smaller than or equal to twice the number of non-zero harmonics taken into consideration.

3.3. The asymptotic covariance matrix of the estimator of the Fourier coefficients of input and output

Let the input u(t) and the response y(t) be periodic with period T. Assume that the observations are described by

$$v(t) = u(t) + g(t) \quad \text{and} \quad z(t) + y(t) + h(t)$$

where g(t) and h(t) are stationary stochastic processes. Let the observation time be JT where J is an integer. Consider the least squares estimator

$$\hat{\eta}^{(J)} = (\hat{\alpha}^{(J)}_{k_1 y}\ \hat{\beta}^{(J)}_{k_1 y}\ \hat{\alpha}^{(J)}_{k_1 u}\ \hat{\beta}^{(J)}_{k_1 u} \ldots . \hat{\alpha}^{(J)}_{k_L y}\ \hat{\beta}^{(J)}_{k_L y}\ \hat{\alpha}^{(J)}_{k_L u}\ \hat{\beta}^{(J)}_{k_L u})' \tag{3.3.1}$$

of the Fourier coefficients

$$\eta = (\alpha_{k_1 y}\ \beta_{k_1 y}\ \alpha_{k_1 u}\ \beta_{k_1 u} \ldots .\ \alpha_{k_L y}\ \beta_{k_L y}\ \alpha_{k_L u}\ \beta_{k_L u})'$$

where $k_1 \ldots k_L$ are the harmonic numbers of the L harmonics taken into consideration. The elements of $\hat{\eta}^{(J)}$ are defined by (3.1.3) and (3.1.4). The asymptotic covariance matrix of $\hat{\eta}^{(J)}$ is defined as

$$\Sigma_{\hat{\eta}\hat{\eta}} = \frac{1}{JT} V_{\hat{\eta}\hat{\eta}} \tag{3.3.2}$$

where $V_{\hat{\eta}\hat{\eta}}$ is defined by

$$V_{\hat{\eta}\hat{\eta}} = \lim_{J\to\infty} E \sqrt{JT}(\hat{\eta}^{(J)} - E\hat{\eta}^{(J)}) \sqrt{JT}(\hat{\eta}^{(J)} - E\hat{\eta}^{(J)})' \tag{3.3.3}$$

The 4L x 4L matrix $V_{\hat{\eta}\hat{\eta}}$ is evaluated by computing all covariances between the elements of $\sqrt{JT}\ \hat{\eta}^{(J)}$. These computations follow closely the procedure for computation of $\lim_{J\to\infty} E(\sqrt{JT}\ \hat{\alpha}^{(J)}_{kf} - E\ \sqrt{JT}\ \hat{\alpha}^{(J)}_{kf})^2$ described in Section 3.1.

Therefore here only the results are given. These may be summarized as follows.

Let $S_{gg}(j\omega)$ and $S_{hh}(j\omega)$ be the power density spectra of $g(t)$ and $h(t)$ respectively and let $S_{gh}(j\omega)$ be the cross power density spectrum of these processes. Assume that the integrals

$$\int_{-\infty}^{\infty} |R_{gg}(t)|dt, \int_{-\infty}^{\infty} |R_{hh}(t)|dt, \int_{-\infty}^{\infty} |R_{gh}(t)|dt$$

$$\int_{-\infty}^{\infty} |t\ R_{gg}(t)|dt, \int_{-\infty}^{\infty} |t\ R_{hh}(t)|dt, \int_{-\infty}^{\infty} |t\ R_{gh}(t)|dt$$

are finite, where $R_{gg}(t)$ and $R_{hh}(t)$ are the autocovariance functions of $g(t)$ and $h(t)$ respectively and $R_{gh}(t)$ is their cross covariance function. Then it can be shown that

$$\lim_{J\to\infty} \operatorname{cov}(\sqrt{JT}\ \hat{\alpha}^{(J)}_{k_\ell u}, \sqrt{JT}\ \hat{\alpha}^{(J)}_{k_m y}) = \begin{cases} \frac{1}{2} \operatorname{Re} S_{gh}(j2\pi k_\ell/T) & k_\ell=k_m \\ 0 & k_\ell\neq k_m \end{cases}$$

$$\lim_{J\to\infty} \operatorname{cov}(\sqrt{JT}\ \hat{\beta}^{(J)}_{k_\ell u}, \sqrt{JT}\ \hat{\beta}^{(J)}_{k_m y}) = \begin{cases} \frac{1}{2} \operatorname{Re} S_{gh}(j2\pi k_\ell/T) & k_\ell=k_m \\ 0 & k_\ell\neq k_m \end{cases}$$

$$\lim_{J\to\infty} \operatorname{cov}(\sqrt{JT}\ \hat{\alpha}^{(J)}_{k_\ell u}, \sqrt{JT}\ \hat{\beta}^{(J)}_{k_m y}) = \begin{cases} \frac{1}{2} \operatorname{Im} S_{gh}(j2\pi k_\ell/T) & k_\ell=k_m \\ 0 & k_\ell\neq k_m \end{cases}$$

$$\lim_{J\to\infty} \operatorname{cov}(\sqrt{JT}\ \hat{\beta}^{(J)}_{k_\ell u}, \sqrt{JT}\ \hat{\alpha}^{(J)}_{k_m y}) = \begin{cases} -\frac{1}{2} \operatorname{Im} S_{gh}(j2\pi k_\ell/T) & k_\ell=k_m \\ 0 & k_\ell\neq k_m \end{cases}$$

(3.3.4)

The elements of $V_{\hat{\eta}\hat{\eta}}$ which are not described by (3.3.4) are obtained by substituting u for y and g for h respectively or by substituting y for u and h for g respectively in (3.3.4).

It follows from the above considerations that

$$V_{\hat{\eta}\hat{\eta}} = \mathrm{diag}\,(W_{k_1} \; \ldots \; W_{k_\ell} \; \ldots \; W_{k_L}) \tag{3.3.5}$$

where

$$W_{k_i} = \frac{1}{2} \begin{pmatrix} S_{hh}(j2\pi k_i/T) & 0 & \mathrm{Re}\, S_{gh}(j2\pi k_i/T) & -\mathrm{Im}\, S_{gh}(j2\pi k_i/T) \\ 0 & S_{hh}(j2\pi k_i/T) & \mathrm{Im}\, S_{gh}(j2\pi k_i/T) & \mathrm{Re}\, S_{gh}(j2\pi k_i/T) \\ \mathrm{Re}\, S_{gh}(j2\pi k_i/T) & \mathrm{Im}\, S_{gh}(j2\pi k_i/T) & S_{gg}(j2\pi k_i/T) & 0 \\ -\mathrm{Im}\, S_{gh}(j2\pi k_i/T) & \mathrm{Re}\, S_{gh}(j2\pi k_i/T) & 0 & S_{gg}(j2\pi k_i/T) \end{pmatrix} \tag{3.3.6}$$

This completes the computation of the asymptotic covariance matrix of $\hat{\eta}^{(J)}$.

3.4. The asymptotic covariance matrix of the estimator of the system parameters

In this section the asymptotic covariance matrix of the estimator of the system parameters is computed. The asymptotic covariance matrix of the residuals is also derived.

The asymptotic covariance matrix $\Sigma_{\hat{\theta}\hat{\theta}}$ of the estimator $\hat{\theta}^{(J)} = (\hat{a}_o^{(J)}\ \hat{a}_1^{(J)}\ \ldots\ \hat{a}_N^{(J)}\ \hat{b}_o^{(J)} \ldots\ \hat{b}_{M-1}^{(J)}\ \hat{\tau}^{(J)})'$ of the system parameters $\theta = (a_o\ \ldots\ a_N\ b_o\ \ldots\ b_{M-1}\ \tau)'$ is defined by

$$\Sigma_{\hat{\theta}\hat{\theta}} = \frac{1}{JT} V_{\hat{\theta}\hat{\theta}} \tag{3.4.1}$$

where

$$V_{\hat{\theta}\hat{\theta}} = \lim_{J\to\infty} E\ \sqrt{JT}(\hat{\theta}^{(J)} - E\hat{\theta}^{(J)})\ \sqrt{JT}(\hat{\theta}^{(J)} - E\hat{\theta}^{(J)})'$$

The elements of $\hat{\theta}^{(J)}$ are functions of the estimator of the Fourier coefficients $\hat{\eta}^{(J)} = (\hat{\alpha}_{k_1y}^{(J)}\ \hat{\beta}_{k_1y}^{(J)}\ \hat{\alpha}_{k_1u}^{(J)}\ \beta_{k_1u}^{(J)}\ \ldots$ $\hat{\alpha}_{k_Ly}^{(J)}\ \hat{\beta}_{k_Ly}^{(J)}\ \hat{\alpha}_{k_Lu}^{(J)}\ \hat{\beta}_{k_Lu}^{(J)})'$ whose asymptotic covariance matrix

$\Sigma_{\hat{\eta}\hat{\eta}}$ was computed in Section 3.3. $\Sigma_{\hat{\theta}\hat{\theta}}$ is computed from $\Sigma_{\hat{\eta}\hat{\eta}}$ using a theorem due to Goldberger. See the Appendix. According to this theorem

$$\Sigma_{\hat{\theta}\hat{\theta}} = \left(\frac{\partial\hat{\theta}^{(J)}}{\partial\hat{\eta}^{(J)}}\right)'_{\hat{\eta}^{(J)}=\eta} \Sigma_{\hat{\eta}\hat{\eta}} \left(\frac{\partial\hat{\theta}^{(J)}}{\partial\hat{\eta}^{(J)}}\right)_{\hat{\eta}^{(J)}=\eta} \tag{3.4.2}$$

where it is assumed that the derivatives exist and that the central moments of the elements of $\hat{\eta}^{(J)}$ of order higher than two are of order of magnitude lower than $\frac{1}{JT}$.

Now recall that the estimator $\hat{\theta}^{(J)}$ is that value of θ which minimizes

$$\hat{\varepsilon}'^{(J)}\ \Omega^{-1}\ \hat{\varepsilon}^{(J)} \tag{3.4.3}$$

where $\hat{\varepsilon}^{(J)} = \hat{P}^{(J)}c - \hat{d}^{(J)}$ denotes the vector of residuals. Thus $\hat{\theta}^{(J)}$ must satisfy

$$\left(\frac{\partial \hat{\varepsilon}^{(J)}}{\partial \theta}\right)_{\theta=\hat{\theta}^{(J)}} \Omega^{-1}(\hat{\varepsilon}^{(J)})_{\theta=\hat{\theta}^{(J)}} = 0 \qquad (3.4.4)$$

This is a set of N+M+2 equations in $\hat{\theta}^{(J)}$ and $\hat{\eta}^{(J)}$. Since at $\hat{\eta}^{(J)} = \eta$ $(\hat{\varepsilon}^{(J)})_{\theta=\hat{\theta}^{(J)}} = 0$, it follows from (3.4.4) that

$$\left\{\left(\frac{\partial \hat{\varepsilon}^{(J)}}{\partial \theta}\right)_{\theta=\hat{\theta}^{(J)}} \Omega^{-1}\left(\frac{\partial(\hat{\varepsilon}^{(J)})_{\theta=\hat{\theta}^{(J)}}}{\partial \hat{\theta}^{(J)}}\right)' d\hat{\theta}^{(J)} + \right.$$

$$\left. + \left(\frac{\partial \hat{\varepsilon}^{(J)}}{\partial \theta}\right)_{\theta=\hat{\theta}^{(J)}} \Omega^{-1}\left(\frac{\partial(\hat{\varepsilon}^{(J)})_{\theta=\hat{\theta}^{(J)}}}{\partial \hat{\eta}^{(J)}}\right)' d\hat{\eta}^{(J)}\right\}_{\hat{\eta}^{(J)}=\eta} = 0$$

and hence

$$\left(\frac{\partial \hat{\theta}^{(J)}}{\partial \hat{\eta}^{(J)}}\right)'_{\hat{\eta}^{(J)}=\eta} =$$

$$- \left\{\frac{\partial \hat{\varepsilon}^{(J)}}{\partial \theta}\Omega^{-1}\left(\frac{\partial \hat{\varepsilon}^{(J)}}{\partial \theta}\right)'\right\}^{-1}_{\hat{\eta}^{(J)}=\eta} \left\{\frac{\partial \hat{\varepsilon}^{(J)}}{\partial \theta}\Omega^{-1}\left(\frac{\partial \hat{\varepsilon}^{(J)}}{\partial \hat{\eta}^{(J)}}\right)'\right\}_{\hat{\eta}^{(J)}=\eta}$$

$$(3.4.5)$$

Defining

$$\Phi' = \left(\frac{\partial \hat{\varepsilon}^{(J)}}{\partial \theta}\right)_{\hat{\eta}^{(J)}=\eta} \qquad (3.4.6)$$

and

$$G' = \left(\frac{\partial \hat{\varepsilon}^{(J)}}{\partial \hat{\eta}^{(J)}}\right)_{\hat{\eta}^{(J)}=\eta} \qquad (3.4.7)$$

(3.4.5) may be written

$$\left(\frac{\partial \hat{\theta}^{(J)}}{\partial \hat{\eta}^{(J)}}\right)'_{\hat{\eta}^{(J)}=\eta} = - (\Phi'\Omega^{-1}\Phi)^{-1}(\Phi'\Omega^{-1}G) \qquad (3.4.8)$$

Combining (3.4.2) and (3.4.8)

$$\Sigma_{\hat{\theta}\hat{\theta}} = (\Phi'\Omega^{-1}\Phi)^{-1}\Phi'\Omega^{-1}G\Sigma_{\hat{\eta}\hat{\eta}}G'\Omega^{-1}\Phi(\Phi'\Omega^{-1}\Phi)^{-1} \qquad (3.4.9)$$

Now consider the matrix $G\Sigma_{\hat{\eta}\hat{\eta}}G'$ in this expression. According to (3.4.7)

$$G\Sigma_{\hat{\eta}\hat{\eta}}G' = \left(\frac{\partial\hat{\varepsilon}^{(J)}}{\partial\hat{\eta}^{(J)}}\right)'_{\hat{\eta}^{(J)}=\eta} \Sigma_{\hat{\eta}\hat{\eta}} \left(\frac{\partial\hat{\varepsilon}^{(J)}}{\partial\hat{\eta}^{(J)}}\right)_{\hat{\eta}^{(J)}=\eta} \qquad (3.4.10)$$

So it follows from Goldberger's theorem that $G\Sigma_{\hat{\eta}\hat{\eta}}G'$ is the asymptotic covariance matrix $\Sigma_{\hat{\varepsilon}\hat{\varepsilon}}$ of the residuals $\hat{\varepsilon}^{(J)}$. Therefore in what follows

$$\Sigma_{\hat{\varepsilon}\hat{\varepsilon}} = G\Sigma_{\hat{\eta}\hat{\eta}}G' \qquad (3.4.11)$$

Computation of $\Sigma_{\hat{\varepsilon}\hat{\varepsilon}}$

In the case of unknown time delay the odd and the even numbered elements of $\hat{\varepsilon}^{(J)}$ are described by

$$\mathrm{Re}\{A(j\omega_{k_\ell})\hat{\gamma}^{(J)}_{k_\ell y} - B_\tau(j\omega_{k_\ell})\hat{\gamma}^{(J)}_{k_\ell u}\} \qquad (3.4.12)$$

and

$$\mathrm{Im}\{A(j\omega_{k_\ell})\hat{\gamma}^{(J)}_{k_\ell y} - B_\tau(j\omega k_\ell)\hat{\gamma}_{k_\ell u}\} \qquad (3.4.13)$$

respectively, where

$$B_\tau(j\omega) = B(j\omega)\exp(-j\omega\tau) \qquad (3.4.14)$$

Then

$$G' = \left(\frac{\partial\hat{\varepsilon}^{(J)}}{\partial\hat{\eta}^{(J)}}\right)_{\hat{\eta}^{(J)}=\eta} = \begin{pmatrix} G'_{k_1} & & 0 \\ & \ddots & \\ 0 & & G'_{k_L} \end{pmatrix} \qquad (3.4.15)$$

where

$$G'_{k_i} = \begin{pmatrix} \mathrm{Re}\ A(j\omega_{k_i}) & \mathrm{Im}\ A(j\omega_{k_i}) \\ \mathrm{Im}\ A(j\omega_{k_i}) & -\mathrm{Re}\ A(j\omega_{k_i}) \\ -\mathrm{Re}\ B_\tau(j\omega_{k_i}) & -\mathrm{Im}\ B_\tau(j\omega_{k_i}) \\ -\mathrm{Im}\ B_\tau(j\omega_{k_i}) & \mathrm{Re}\ B\ (j\omega_{k_i}) \end{pmatrix} \qquad (3.4.16)$$

The matrix $\Sigma_{\hat{\eta}\hat{\eta}}$ in (3.4.11) has been computed in Section 3.3 and is described by (3.3.2), (3.3.5) and (3.3.6). Inserting (3.3.4) and (3.4.15) in (3.4.11) and multiplying

$$\Sigma_{\hat{\varepsilon}\hat{\varepsilon}} = \frac{1}{JT}\ \mathrm{diag}\ (D_{k_1}, D_{k_2}, \ldots, D_{k_L}) \qquad (3.4.17)$$

where the 2 x 2 matrices D_ℓ are described by

$$D_{k_\ell} = G_{k_\ell} W_{k_\ell} G'_{k_\ell} \qquad (3.4.18)$$

Then using (3.3.6) and (3.4.16) it is easily shown that the off-diagonal elements of D_{k_L} are zero, while the diagonal elements are described by

$$\frac{1}{2}|A(j\omega_{k_\ell})|^2 S_{hh}(j\omega_{k_\ell}) + \frac{1}{2}|B(j\omega_{k_\ell})|^2 S_{gg}(j\omega_{k_\ell}) +$$

$$+\{-\mathrm{Re}\ A(j\omega_{k_\ell})\mathrm{Re}\ B_\tau(j\omega_{k_\ell}) +$$

$$-\ \mathrm{Im}\ A(j\omega_{k_\ell})\mathrm{Im}\ B_\tau(j\omega_{k_\ell})\}\mathrm{Re}\ S_{gh}(j\omega_{k_\ell}) +$$

$$+\{-\mathrm{Im}\ A(j\omega_{k_\ell})\mathrm{Re}\ B_\tau(j\omega_{k_\ell}) +$$

$$+\ \mathrm{Re}\ A(j\omega_{k_\ell})\mathrm{Im}\ B\ (j\omega_{k_\ell})\}\mathrm{Im}\ S_{gh}(j\omega_{k_\ell}) =$$

$$= \frac{1}{2}\,\mathrm{Re}\{|A(j\omega_{k_\ell})|^2 S_{hh}(j\omega_{k_\ell}) + |B_\tau(j\omega_{k_\ell})|^2 S_{gg}(j\omega_{k_\ell}) +$$

$$+ 2A(j\omega_{k_\ell})B_\tau^{\star}(j\omega_{k_\ell})S_{gh}(j\omega_{k_\ell})\} \qquad (3.4.19)$$

Now let the processes $g(t)$ and $h(t)$ be composed as described in Section 1.2, that is

$$h(t) = h_1(t) + h_2(t) + h_3(t)$$

and

$$g(t) = g_1(t) + g_2(t)$$

where $h_1(t)$ is the steady state response of the system to $g_1(t)$, $g_2(t)$ and $h_2(t)$ are mutually independent and independent of all other components of $g(t)$ and $h(t)$; $h_3(t)$ is possibly correlated with both $g_1(t)$ and $h_1(t)$. Under these assumptions $S_{hh}(j\omega)$, $S_{gg}(j\omega)$ and $S_{gh}(j\omega)$ may be decomposed as follows

$$S_{hh}(j\omega) = S_{h_1h_1}(j\omega) + S_{h_2h_2}(j\omega) + S_{h_3h_3}(j\omega) +$$

$$+ 2\,\mathrm{Re}\,S_{h_1h_3}(j\omega) \qquad (3.4.20)$$

$$S_{gg}(j\omega) = S_{g_1g_1}(j\omega) + S_{g_2g_2}(j\omega) \qquad (3.4.21)$$

and

$$S_{gh}(j\omega) = S_{g_1h_1}(j\omega) + S_{g_1h_3}(j\omega) \qquad (3.4.22)$$

Furthermore, since $h_1(t)$ is the steady state response to $g_1(t)$

$$|A(j\omega)|^2 S_{h_1h_1}(j\omega) + |B_\tau(j\omega)|^2 S_{g_1g_1}(j\omega) +$$

$$- 2A(j\omega)B_\tau^{\star}(j\omega)S_{g_1h_1}(j\omega) =$$

$$= |A(j\omega)|^2 S_{h_1h_1}(j\omega) + |B_\tau(j\omega)|^2 S_{g_1g_1}(j\omega) +$$

$$- 2A(j\omega)B_\tau^{\star}(j\omega)\{B_\tau(j\omega)/A(j\omega)\}S_{g_1g_1}(j\omega)$$

$$= |A(j\omega)|^2 S_{h_1h_1}(j\omega) - |B_\tau(j\omega)|^2 S_{g_1g_1}(j\omega) = 0 \tag{3.4.23}$$

and

$$S_{h_1h_3}(j\omega) = \{B_\tau^{\star}(j\omega)/A^{\star}(j\omega)\}S_{g_1h_3}(j\omega) \tag{3.4.24}$$

Then using (3.4.19) and (3.4.20) - (3.4.24) it is easily shown that the diagonal elements of D_{k_ℓ} are described by

$$\frac{1}{2}[\,|A(j\omega_{k_\ell})|^2\{S_{h_2h_2}(j\omega_{k_\ell}) + S_{h_3h_3}(j\omega_{k_\ell})\} + \\ + |B_\tau(j\omega_{k_\ell})|^2 S_{g_2g_2}(j\omega_{k_\ell})]$$

It then follows from (3.4.17) that

$$\Sigma_{\hat{\varepsilon}\hat{\varepsilon}} = \frac{1}{JT}\,\mathrm{diag}\{\lambda(\omega_{k_1}), \lambda(\omega_{k_1}), \ldots, \lambda(\omega_{k_L}), \lambda(\omega_{k_L})\} \tag{3.4.25}$$

where

$$\lambda(\omega_{k_\ell}) = \frac{1}{2}[\,|A(j\omega_{k_\ell})|^2\{S_{h_2h_2}(j\omega_{k_\ell}) + S_{h_3h_3}(j\omega_{k_\ell})\} + \\ + |B_\tau(j\omega_{k_\ell})|^2 S_{g_2g_2}(j\omega_{k_\ell})] \tag{3.4.26}$$

This completes the computation of $\Sigma_{\hat{\varepsilon}\hat{\varepsilon}}$.

Finally, denoting

$$\Lambda = \mathrm{diag}\{\lambda(\omega_{k_1}), \lambda(\omega_{k_1}), \ldots, \lambda(\omega_{k_L}), \lambda(\omega_{k_L})\} \tag{3.4.27}$$

(3.4.26) may be written

$$\Sigma_{\hat{\varepsilon}\hat{\varepsilon}} = \frac{1}{JT} \Lambda \tag{3.4.28}$$

Hence it follows from (3.4.9) and (3.4.11) that

$$\Sigma_{\hat{\theta}\hat{\theta}} = \frac{1}{JT} (\Phi'\Omega^{-1}\Phi)^{-1}\Phi'\Omega^{-1}\Lambda\Omega^{-1}\Phi(\Phi'\Omega^{-1}\Phi)^{-1} \tag{3.4.29}$$

Computation of $\Sigma_{\tilde{\theta}\tilde{\theta}}$

If $\Omega = \Lambda$ the corresponding estimator is the minimum variance least squares estimator $\tilde{\theta}$ described in Section 1.2. It follows from (3.4.29) that

$$\Sigma_{\tilde{\theta}\tilde{\theta}} = \frac{1}{JT} (\Phi'\Lambda^{-1}\Phi)^{-1} \tag{3.4.30}$$

The elements of $\Sigma_{\tilde{\theta}\tilde{\theta}}$ are computed as follows. According to (3.4.6)

$$\Phi' = \left(\frac{\partial \hat{\varepsilon}^{(J)}}{\partial \theta}\right)_{\hat{\eta}^{(J)}=\eta} \tag{3.4.31}$$

where the elements of $\hat{\varepsilon}^{(J)}$ are defined by (3.4.12)-(3.4.14). Differentiating $\hat{\varepsilon}^{(J)}$ with respect to $\theta = (a_o, \ldots, a_N, b_o, \ldots b_{M-1}\tau)'$ yields

$$\Phi = (\Phi'_{k_1} \; \Phi'_{k_2} \; \ldots \; \Phi'_{k_L})' \tag{3.4.32}$$

where

$$\Phi_{k_\ell} = \begin{pmatrix} \text{Re } s_{k_\ell}^{o}\gamma_{k_\ell}y & \ldots & \text{Re } s_{k_\ell}^{N}\gamma_{k_\ell}y & -\text{Re } s_{k_\ell}^{o}\gamma_{k_\ell}u\exp(-s_{k_\ell}\tau) & \ldots & -\text{Re } s_{k_\ell}^{M-1}\gamma_{k_\ell}u \exp(-s_{k_\ell}\tau) & -\text{Re } -s_{k_\ell}\gamma_{k_\ell}u^{B}{}_{\tau}(s_{k_\ell}) \\ \text{Im } s_{k_\ell}^{o}\gamma_{k_\ell}y & \ldots & \text{Im } s_{k_\ell}^{N}\gamma_{k_\ell}y & -\text{Im } s_{k_\ell}^{o}\gamma_{k_\ell}u\exp(-s_{k_\ell}\tau) & \ldots & -\text{Im } s_{k_\ell}^{M-1}\gamma_{k_\ell}u \exp(-s_{k_\ell}\tau) & -\text{Im } -s_{k_\ell}\gamma_{k_\ell}u^{B}{}_{\tau}(s_{k_\ell}) \end{pmatrix} \tag{3.4.33}$$

with $s_{k_\ell} = j\omega_{k_\ell}$. Now define

$$S_{uu}(j\omega_{k_\ell}) = |\gamma_{k_\ell u}|^2$$

Furthermore note that

$$\gamma_{k_\ell y} = \{B_\tau(j\omega_{k_\ell})/A(j\omega_{k_\ell})\}\gamma_{k_\ell u}$$

Then using (3.4.27), (3.4.32) and (3.4.33) it is easily shown that the elements x_{pq} of the matrix $X = JT(\Phi'\Lambda^{-1}\Phi)$ may be written

$$x_{pq} = JT \sum_{l=1}^{L} \frac{S_{uu}(j\omega_{k_\ell})}{\lambda(\omega_{k_\ell})} \mathrm{Re}\{F_p(j\omega_{k_\ell})F_q^{*}(j\omega_{k_\ell})\} \qquad (3.4.34)$$

where

$$F_i(j\omega) = (j\omega)^{i-1}B_\tau(j\omega)/A(j\omega) \qquad i = 1, \ldots, N+1$$

$$F_i(j\omega) = -(j\omega)^{i-N-2}\exp(-j\omega\tau) \qquad i = N+2, \ldots, N+M+1$$

and

$$F_i(j\omega) = j\omega\, B_\tau(j\omega) \qquad i = N+M+2$$

Next consider the special case that

$g(t) = g_1(t) = g_2(t) = 0.$

This implies that $h_1(t) = 0$ since $h_1(t)$ is the steady state response to $g_1(t)$. Now from (3.4.20) and (3.4.21)

$$S_{hh}(j\omega) = S_{h_1h_1}(j\omega) + S_{h_2h_2}(j\omega)$$

and

$$S_{gg}(j\omega) = 0.$$

Hence in view of (3.4.26)

$$\lambda(\omega) = \tfrac{1}{2}|A(j\omega)|^2\, S_{hh}(j\omega)$$

Inserting this in (3.4.34) gives

$$x_{pq} = JT \sum_{l=1}^{L} 2 \frac{S_{uu}(j\omega_{k_\ell})}{S_{hh}(j\omega_{k_\ell})|A(j\omega_{k_\ell})|^2} \mathrm{Re}\{F_p(j\omega_{k_\ell})F_q^{*}(j\omega_{k_\ell})\}$$

$$= JT \sum_{\substack{l=-L \\ l\neq 0}}^{L} \frac{S_{uu}(j\omega_{k_\ell})}{S_{hh}(j\omega_{k_\ell})|A(j\omega_{k_\ell})|^2} \mathrm{Re}\{F_p(j\omega_{k_\ell})F_q^{*}(j\omega_{k_\ell})\}$$

(3.4.35)

Defining

$$H_i(j\omega) = \frac{(j\omega)^{i-1}}{A(j\omega)} \qquad i = 1, \ldots, N+1$$

$$H_i(j\omega) = -\frac{(j\omega)^{i-N-2}}{B(j\omega)} \qquad i = N+2, \ldots, N+M+1 \qquad (3.4.36)$$

and

$$H_i(j\omega) = j\omega \qquad i = N+M+2$$

(3.4.35) may be written

$$x_{pq} = JT \sum_{\substack{l=-L \\ l\neq 0}}^{L} \frac{S_{uu}(j\omega_{k_\ell})}{S_{hh}(j\omega_{k_\ell})} |H_S(j\omega_{k_\ell})|^2 \mathrm{Re}\{H_p(j\omega_{k_\ell})H_q^{*}(j\omega_{k_\ell})\}$$

(3.4.37)

where

$$H_S(j\omega) = B_\tau(j\omega)/A(j\omega).$$

So the asymptotic covariance matrix $\Sigma_{\tilde{\theta}\tilde{\theta}}$ of the minimum variance least squares estimator $\tilde{\theta}$ for the case that $g(t) = 0$ is obtained by inverting the matrix X whose elements are defined by (3.4.36) and (3.4.37). In Section 3.5 it will be shown that in this particular case $\Sigma_{\tilde{\theta}\tilde{\theta}}$ is equal to the minimum variance bound for the estimation of θ if $h(t)$ is normally distributed.

3.5. The minimum variance bound

This section computes the minimum variance bound (MVB) for the estimation of the coefficients of the differential equation and the time delay of the syst. ..

Let the system be described by

$$a_N \frac{d^N y(t)}{dt^N} + a_{N-1} \frac{d^{N-1} y(t)}{dt^{N-1}} + \ldots + a_o y(t) =$$

$$= \frac{d^M u(t-\tau)}{dt^M} + b_{M-1} \frac{d^{M-1} u(t-\tau)}{dt^{M-1}} + \ldots + b_o u(t-\tau) \qquad (3.5.1)$$

where $u(t)$ is the input, $y(t)$ is the response to $u(t)$ and τ denotes time delay. The transfer function of the system (3.5.1) is given by

$$H_S(j\omega) = \frac{B(j\omega)}{A(j\omega)}\exp(-j\omega\tau) \qquad (3.5.2)$$

with

$$A(j\omega) = a_N(j\omega)^N + a_{N-1}(j\omega)^{N-1} + \ldots + a_o \qquad (3.5.3)$$

and

$$B(j\omega) = (j\omega)^M + b_{M-1}(j\omega)^{M-1} + \qquad + b_o \qquad (3.5.4)$$

where ω is frequency in rad s^{-1}.

The response $y(t)$ may be expressed by the convolution integral

$$y(t) = \int_o^\infty h_S(t_1)u(t-t_1)dt_1 \qquad (3.5.5)$$

where $h_S(t)$ denotes the impulse response of the system. Defining

$$\theta_S = (a_o \ldots a_N\ b_o \ldots b_{M-1}\ \tau)' \qquad (3.5.6)$$

it then follows from (3.5.5) that

$$\frac{\partial y(t)}{\partial \theta_S} = \frac{\partial}{\partial \theta_S} \int_o^\infty h_S(t_1)u(t-t_1)t_1$$

$$= \int_o^\infty \frac{\partial h_S(t_1)}{\partial \theta_S} u(t-t_1)t_1 \qquad (3.5.7)$$

The changing of order of integrating and differentiating is allowed if $h_S(t_1)$ is continuously differentiable with respect to θ_S. It is concluded from (3.5.7) that $\frac{\partial y(t)}{\partial \theta_S}$ is the response to u(t) of a system having impulse response $\frac{\partial h_S(t)}{\partial \theta_S}$. Also $h_S(t)$ and $H_S(j\omega)$ form a Fourier transform pair. Therefore

$$H_S(j\omega) = \int_{-\infty}^{\infty} h_S(t)\exp(-j\omega t)\,dt$$

and hence

$$\frac{\partial H_S(j\omega)}{\partial \theta_S} = \int_{-\infty}^{\infty} \frac{\partial h_S(t)}{\partial \theta_S}\exp(-j\omega t)\,dt$$

So $\frac{\partial y(t)}{\partial \theta_S}$ is the response to u(t) of a system having transfer function $\frac{\partial H_S(j\omega)}{\partial \theta_S}$. In what follows the expression

$$y(t) = \mathscr{F}^{-1}\{H_S(j\omega)\} \star u(t) \qquad (3.5.8)$$

will indicate that y(t) is the response of the system to u(t); $\mathscr{F}^{-1}$ and $\star$ refer to the inverse Fourier transform and the convolution respectively. It then follows from the above considerations that

$$\frac{\partial y(t)}{\partial \theta_S} = \mathscr{F}^{-1}\left\{\frac{\partial H_S(j\omega)}{\partial \theta_S}\right\} \star u(t)$$

Now assume that

$$z(t) = y(t) + h(t) \qquad (3.5.9)$$

is observed at the output of the system. The disturbance h(t) is assumed to be a stationary, normally distributed process described by

$$h(t) = \mathscr{F}^{-1}\{H_D(j\omega)\} \star e(t) \qquad (3.5.10)$$

where $H_D(j\omega)$ is a transfer function and $e(t)$ is a zero mean, stationary, normally distributed process having autocovariance function

$$E\, e(t)\, e(t') = \delta(t-t')$$

Furthermore it is assumed that $u(t)$ and $e(t)$ are independent and that $H_D(j\omega)$ has all zero's in the upper half-plane.

Now let $u(t)$ and $z(t)$ be observed for $0 < t < T_{obs}$ and let the vector θ be defined by

$$\theta = (\theta_S{}'\theta_D{}')'$$

where θ_S is defined by (3.5.6) and θ_D is the vector of the parameters of $H_D(j\omega)$. The elements of θ_D will be referred to as noise parameters. Then using a procedure due to Åström (1967) the MVB for the estimation of θ is computed as follows. First it is observed that the logarithm of the likelihood function L of the normally distributed $e(t)$ for $0 < t < T_{obs}$ may be written

$$\ln L = -\tfrac{1}{2}\int_0^{T_{obs}} e^2(t)dt + \text{constant} \qquad (3.5.11)$$

See Åström (1967). Generally the relation between the logarithm of the likelihood function and the MVB of any unbiased estimator $\check{\theta}$ of the parameters of the likelihood function is

$$E(\check{\theta} - \theta)(\check{\theta} - \theta)' \geqslant \Pi^{-1}$$

where $\Pi = [\,\pi_{pq}]$ is the information matrix defined by

$$\pi_{pq} = - E\,\frac{\partial^2 \ln L}{\partial\theta_p \partial\theta_q} \qquad (3.5.12)$$

See Kendall and Stuart (1967), page 53. Now it follows from (3.5.11) and (3.5.12) that

$$\pi_{pq} = E\int_0^{T_{obs}} \frac{\partial e(t)}{\partial \theta_p}\,\frac{\partial e(t)}{\partial \theta_q}\,dt + E\int_0^{T_{obs}} e(t)\frac{\partial^2 e(t)}{\partial \theta_p \partial \theta_q}\,dt \tag{3.5.13}$$

Combining (3.5.8), (3.5.9) and (3.5.10) and rearranging

$$\mathscr{F}^{-1}\{H_D(j\omega)\}\star e(t) = z(t) - \mathscr{F}^{-1}\{H_S(j\omega)\}\star u(t)$$

Hence

$$\mathscr{F}^{-1}\{H_D(j\omega)\}\star\ \frac{\partial e(t)}{\partial \theta_S} = -\,\mathscr{F}^{-1}\left\{\frac{\partial H_S(j\omega)}{\partial \theta_S}\right\}\star u(t) \tag{3.5.14}$$

and

$$\mathscr{F}^{-1}\{H_D(j\omega)\}\star\ \frac{\partial e(t)}{\partial \theta_D} = -\,\mathscr{F}^{-1}\left\{\frac{\partial H_D(j\omega)}{\partial \theta_D}\right\}\star e(t) \tag{3.5.15}$$

Furthermore

$$\mathscr{F}^{-1}\{H_D(j\omega)\}\star\ \frac{\partial^2 e(t)}{\partial \theta_S \partial \theta_D} = -\mathscr{F}^{-1}\left\{\frac{\partial H_D(j\omega)}{\partial \theta_D}\right\}\star\frac{\partial e(t)}{\partial \theta_S}$$

$$= \mathscr{F}^{-1}\left\{\frac{\partial H_D(j\omega)}{\partial \theta_D}\,\frac{1}{H_D(j\omega)}\,\frac{\partial H_S(j\omega)}{\partial \theta_S}\right\}\star u(t) \tag{3.5.16}$$

Since by assumption u(t) and e(t) are independent it follows from (3.5.12) - (3.5.16) that

$$E\,\frac{\partial^2 \ell n\, L}{\partial \theta_S \partial \theta_D} = 0$$

Hence Π may be written

$$\Pi = -\begin{pmatrix} E\,\dfrac{\partial^2 \ell n\, L}{\partial {\theta_S}^2} & 0 \\ 0 & E\,\dfrac{\partial^2 \ell n\, L}{\partial {\theta_D}^2} \end{pmatrix}$$

Consequently the MVB is

$$\Pi^{-1} = - \begin{pmatrix} \left(E \dfrac{\partial^2 \ln L}{\partial \theta_S^2}\right)^{-1} & 0 \\ 0 & \left(E \dfrac{\partial^2 \ln L}{\partial \theta_D^2}\right)^{-1} \end{pmatrix} \tag{3.5.17}$$

From this expression it follows that, for the model of system and noise assumed here, those elements of Π^{-1} which represent the MVB of the system parameter estimates do not depend on whether or not the noise parameters are known. Furthermore it follows from (3.5.13) and (3.5.15) that those elements of Π^{-1} which represent the MVB of the noise parameter estimates do not depend on u(t).

In what follows only the MVB for the estimation of the system parameters will be considered. The expression (3.5.17) shows that this MVB is described by

$$\Psi^{-1} = - \left(E \frac{\partial^2 \ln L}{\partial \theta_S^2}\right)^{-1} \tag{3.5.18}$$

In view of (3.5.12) and (3.5.13)

$$\Psi = E \int_0^{T_{obs}} \frac{\partial e(t)}{\partial \theta_S} \left(\frac{\partial e(t)}{\partial \theta_S}\right)' dt +$$

$$+ E \int_0^{T_{obs}} e(t) \frac{\partial^2 e(t)}{\partial \theta_S^2} dt \tag{3.5.19}$$

Since by assumption e(t) and u(t) are independent it follows from (3.5.14) that the second term of (3.5.19) is zero and hence

$$\Psi = E \int_0^{T_{obs}} \frac{\partial e(t)}{\partial \theta_S} \left(\frac{\partial e(t)}{\partial \theta_S}\right)' dt = T_{obs}\, E \frac{\partial e(t)}{\partial \theta_S} \left(\frac{\partial e(t)}{\partial \theta_S}\right)' \tag{3.5.20}$$

The elements of the vector $\frac{\partial e(t)}{\partial \theta_S}$ in this expression are computed from (3.5.14) and (3.5.2). These elements are

$$\frac{\partial e(t)}{\partial a_i} = \mathscr{F}^{-1}\left\{\frac{(j\omega)^i\ H_S(j\omega)}{H_D(j\omega)A(j\omega)}\right\}\star u(t) \tag{3.5.21}$$

$$\frac{\partial e(t)}{\partial b_i} = -\mathscr{F}^{-1}\left\{\frac{(j\omega)^i\ H_S(j\omega)}{H_D(j\omega)B(j\omega)}\right\}\star u(t) \tag{3.5.22}$$

and

$$\frac{\partial e(t)}{\partial \tau} = \mathscr{F}^{-1}\left\{\frac{j\omega\ H_S(j\omega)}{H_D(j\omega)}\right\}\star u(t) \tag{3.5.23}$$

Periodic test signals

It is observed that (3.5.21) - (3.5.23) are of the form $\mathscr{F}^{-1}\{F(j\omega)\}\star u(t)$. Now assume that $u(t)$ is periodic with period T. Furthermore let $T_{obs} = JT$ where J is an integer. It is easily shown that for two periodic signals $\mathscr{F}^{-1}\{F_p(j\omega)\}\star u(t)$ and $\mathscr{F}^{-1}\{F_q(j\omega)\}\star u(t)$

$$\int_0^{JT} \mathscr{F}^{-1}\{F_p(j\omega)\}\star u(t)\ \mathscr{F}^{-1}\{F_q(j\omega)\}\star u(t)dt$$

$$= JT \sum_{\ell=-\infty}^{\infty} \mathrm{Re}\{F_p(j\omega_\ell)F_q^\star(j\omega_\ell)\}S_{uu}(j\omega_\ell) \tag{3.5.24}$$

where $\omega_\ell = 2\pi\ell/T$ and $S_{uu}(j\omega_\ell) = |\gamma_{\ell u}|^2$. It then follows from (3.5.21) - (3.5.24) that

$$\psi_{pq} = JT \sum_{\ell=-\infty}^{\infty} \frac{S_{uu}(j\omega_\ell)}{S_{hh}(j\omega_\ell)}|H_S(j\omega_\ell)|^2\mathrm{Re}\{H_p(j\omega_\ell)H_q^\star(j\omega_\ell)\} \tag{3.5.25}$$

where $S_{hh}(j\omega) = |H_D(j\omega)|^2$ denotes the power density spectrum of h(t) and

$$H_i(j\omega) = \frac{(j\omega)^{i-1}}{A(j\omega)} \qquad \text{for } i = 1, \ldots, N+1 \qquad (3.5.26)$$

$$H_i(j\omega) = -\frac{(j\omega)^{i-N-2}}{B(j\omega)} \qquad \text{for } i = N+2, \ldots , N+M+1 \qquad (3.5.27)$$

and

$$H_i(j\omega) = j\omega \qquad \text{for } i = N+M+2 \qquad (3.5.28)$$

This completes the computation of the elements of the matrix ψ for the case of periodic inputs. The MVB for the estimation of the system parameters is obtained by inversion of Ψ.

A comparison of the elements ψ_{pq} defined by (3.5.25) - (3.5.28) with the elements x_{pq} defined by (3.4.36) and (3.4.37) shows that the matrix Ψ and the matrix X are identical. Hence Ψ^{-1} and X^{-1} are identical. Since X^{-1} is the asymptotic covariance matrix of the minimum variance least squares estimator $\tilde{\theta}$, it is concluded that $\tilde{\theta}$ asymptotically achieves the MVB if the noise is normally distributed.

3.6. Estimation of the weighting matrix for the two-step least squares procedure

In the two-step least squares procedure described in Section 1.3 the diagonal elements of the matrix Λ are estimated by

$$\hat{\lambda}_w(\omega_{k_i}) = \frac{JT}{2(2k_w+1)} \sum_{k=k_iJ-k_w}^{k_iJ+k_w} \mathrm{Re}^2\{A(j2\pi k/JT)Z(k) +$$

$$- B(j2\pi k/JT)\exp(-j2\pi k\tau/JT)V(k)\}_{\theta=\hat{\theta}_I^{(J)}} +$$

$$+ \mathrm{Im}^2\{A(j2\pi k/JT)Z(k) +$$

$$- B(j2\pi k/JT)\exp(-j2\pi k\tau/JT)V(k)\}_{\theta=\hat{\theta}_I^{(J)}} \qquad (3.6.1)$$

The selection of this estimator was based on the assumptions that asymptotically the square of the measured residual

$$\mathrm{Re}\{A(j2\pi k/JT)Z(k) - B(j2\pi k/JT)\exp(-j2\pi k\tau/JT)V(k)\}_{\theta=\theta_I^{(J)}}$$

and the square of the measured residual

$$\mathrm{Im}\{A(j2\pi k/JT)Z(k) - B(j2\pi k/JT)\exp(-j2\pi k\tau/JT)V(k)\}_{\theta=\hat{\theta}_I^{(J)}}$$

are unbiased estimators of the asymptotic variance of the true residual

$$\mathrm{Re}\{A(j2\pi k/JT)Z(k) - B(j2\pi k/JT)\exp(-j2\pi k\tau/JT)V(k)\}$$

and the asymptotic variance of the true residual

$$\mathrm{Im}\ A(j2\pi k/JT)Z(k) - B(j2\pi k/JT)\exp(-j2\pi k\tau/JT)V(k)\}$$

respectively and that these variances are equal. This section investigates the conditions under which these assumptions hold. First it is shown that the asymptotic expectation of the measured residuals is zero. Next the asymptotic covariance matrix of the measured residuals is computed. Since the asymptotic expectation of the measured residuals is zero, the diagonal elements of this covariance matrix are the asymptotic expectations of the squares of the measured residuals. These diagonal elements are then compared to the corresponding elements of the asymptotic covariance matrix of the true residuals. From this comparison it follows that the above assumption is valid only for those residuals which do not correspond to frequencies of non-zero harmonics of the test signal. This observation results in a proposal for a slight modification of $\hat{\lambda}_w(\omega_{k_i})$.

In the remainder of this section it is assumed that in (3.6.1) $k_i = k_1$ since this simplifies the expressions to be derived. This is no loss of generality since k_1 may refer to any harmonic. Furthermore define

$$s_{k/J} = j2\pi k/JT \quad \text{and} \quad V_\tau(k) = V(k)\exp(-s_{k/J}\tau)$$

and recall that according to (1.3.1) and (1.3.2)

$$V(k) = \frac{1}{JT}\int_0^{JT} v(t)\exp(-j2\pi kt/JT)\,dt \tag{1.3.1}$$

and

$$Z(k) = \frac{1}{JT}\int_0^{JT} z(t)\exp(-j2\pi kt/JT)\,dt \tag{1.3.2}$$

Define the matrix $\hat{P}^{(J)}_{k/J,w}$ by

$$\hat{P}^{(J)}_{k/J,w} = \left(\begin{matrix} \mathrm{Re}\, s^0_{k/J} Z(k) & \mathrm{Re}\, s^1_{k/J} Z(k) & \ldots & \mathrm{Re}\, s^N_{k/J} Z(k) \\ \mathrm{Im}\, s^0_{k/J} Z(k) & \mathrm{Im}\, s^1_{k/J} Z(k) & \ldots & \mathrm{Im}\, s^N_{k/J} Z(k) \end{matrix}\right.$$

$$\left.\begin{matrix} -\mathrm{Re}\, s^0_{k/J} V_\tau(k) & -\mathrm{Re}\, s^1_{k/J} V_\tau(k) & \ldots & -\mathrm{Re}\, s^{M-1}_{k/J} V_\tau(k) \\ -\mathrm{Im}\, s^0_{k/J} V_\tau(k) & -\mathrm{Im}\, s^1_{k/J} V_\tau(k) & \ldots & -\mathrm{Im}\, s^{M-1}_{k/J} V_\tau(k) \end{matrix}\right)$$

(3.6.2.)

and let

$$\hat{P}_w^{(J)} = (\hat{P}'^{(J)}_{(k_1J-k_w)/J,w} \cdots$$

$$\cdots \hat{P}'^{(J)}_{(k_1J-1)/J,w} \; \hat{P}'^{(J)}_{k_1J/J,w} \; \hat{P}'^{(J)}_{(k_1J+1)/J,w} \cdots$$

$$\cdots \hat{P}'^{(J)}_{(k_1J+k_w)/J,w})' \tag{3.6.3}$$

while the vector $\hat{d}_w^{(J)}$ is defined by

$$\hat{d}_w^{(J)} =$$

$$\Big(\mathrm{Re}\; s^M_{(k_1J-k_w)/J}\, V_\tau(k_1J-k_w) \quad \mathrm{Im}\; s^M_{(k_1J-k_w)/J}\, V_\tau(k_1J-k_w)$$

$$\mathrm{Re}\; s^M_{(k_1J-k_w+1)/J}\, V_\tau(k_1J-k_w+1) \quad \mathrm{Im}\; s^M_{(k_1J-k_w+1)/J}\, V_\tau(k_1J-k_w+1)$$

$$\cdots \quad \mathrm{Re}\; s^M_{(k_1J-1)/J}\, V_\tau(k_1J-1) \quad \mathrm{Im}\; s^M_{(k_1J-1)/J}\, V_\tau(k_1J-1)$$

$$\mathrm{Re}\; s^M_{k_1J/J}\, V_\tau(k_1J) \quad \mathrm{Im}\; s^M_{k_1J/J}\, V_\tau(k_1J)$$

$$\mathrm{Re}\; s^M_{(k_1J+1)/J}\, V_\tau(k_1J+1) \quad \mathrm{Im}\; s^M_{(k_1J+1)/J}\, V_\tau(k_1J+1) \cdots$$

$$\mathrm{Re}\; s^M_{(k_1J+k_w-1)/J}\, V_\tau(k_1J+k_w-1) \quad \mathrm{Im}\; s^M_{(k_1J+k_w-1)/J}\, V_\tau(k_1J+k_w-1)$$

$$\mathrm{Re}\; s^M_{(k_1J+k_w)/J}\, V_\tau(k_1J+k_w) \quad \mathrm{Im}\; s^M_{(k_1J+k_w)/J}\, V_\tau(k_1J+k_w)\Big)' \tag{3.6.4}$$

Then the vector of residuals $\hat{\varepsilon}_w^{(J)}$ is defined by

$$\hat{\varepsilon}_w^{(J)} = \hat{P}_w^{(J)}\, c - \hat{d}_w^{(J)}$$

while the vector of measured residuals $\hat{\varepsilon}_{w\hat{\theta}}^{(J)}$ is defined by

$$\hat{\varepsilon}_{w\hat{\theta}}^{(J)} = \left(\hat{P}_w^{(J)} \; c - \hat{d}_w^{(J)}\right)_{\theta=\hat{\theta}_I^{(J)}} \tag{3.6.5}$$

It is observed that the elements of $\hat{\varepsilon}_{w\hat{\theta}}^{(J)}$ are functions of the estimator of the system parameters $\hat{\theta}_I^{(J)}$ and of the elements of the vector $\hat{\eta}_w^{(J)}$ defined by

$$\hat{\eta}_w^{(J)} =$$

$$\begin{aligned}
&(\text{Re } Z(k_1 J-k_w) \;\; -\text{Im } Z(k_1 J-k_w) \;\; \text{Re } V(k_1 J-k_w) \;\; -\text{Im } V(k_1 J-k_w) \\
&\ldots \;\; \text{Re } Z(k_1 J-1) \;\; -\text{Im } Z(k_1 J-1) \;\; \text{Re } V(k_1 J-1) \;\; -\text{Im } V(k_1 J-1) \\
&\quad \text{Re } Z(k_1 J+1) \;\; -\text{Im } Z(k_1 J+1) \;\; \text{Re } V(k_1 J+1) \;\; -\text{Im } V(k_1 J+1) \\
&\ldots \;\; \text{Re } Z(k_1 J+k_w) \;\; -\text{Im } Z(k_1 J+k_w) \;\; \text{Re } V(k_1 J+k_w) \;\; -\text{Im } V(k_1 J+k_w) \\
&\quad \text{Re } Z(k_1 J) \;\; -\text{Im } Z(k_1 J) \;\; \text{Re } V(k_1 J) \;\; -\text{Im } V(k_1 J)\;)'
\end{aligned}$$

Note that the last four elements of $\hat{\eta}_w^{(J)}$ are identical to $\hat{\alpha}_{k_1 y}^{(J)}$, $\hat{\beta}_{k_1 y}^{(J)}$, $\hat{\alpha}_{k_1 u}^{(J)}$ and $\hat{\beta}_{k_1 u}^{(J)}$ respectively. The elements of $\hat{\theta}_I^{(J)}$ are functions of the estimator $\hat{\eta}^{(J)}$ of the Fourier coefficients of input and output defined by (3.3.1). This expression is repeated here

$$\hat{\eta}^{(J)} = (\hat{\alpha}_{k_1 y}^{(J)} \; \hat{\beta}_{k_1 y}^{(J)} \; \hat{\alpha}_{k_1 u}^{(J)} \; \hat{\beta}_{k_1 u}^{(J)} \ldots$$

$$\ldots \; \hat{\alpha}_{k_L y}^{(J)} \; \hat{\beta}_{k_L y}^{(J)} \; \hat{\alpha}_{k_L u}^{(J)} \; \hat{\beta}_{k_L u}^{(J)} \;)' \tag{3.3.1}$$

The elements of $\hat{\eta}_w^{(J)}$ and $\hat{\eta}^{(J)}$ can be combined into a single vector $\hat{\eta}_c^{(J)}$ defined by

$\hat{\eta}_c^{(J)} =$

$$(\mathrm{Re}\ Z(k_1 J-k_w) \quad -\mathrm{Im}\ Z(k_1 J-k_w) \quad \mathrm{Re}\ V(k_1 J-k_w) \quad -\mathrm{Im}\ V(k_1 J-k_w)$$

$$\ldots\ \mathrm{Re}\ Z(k_1 J-1) \quad -\mathrm{Im}\ Z(k_1 J-1) \quad \mathrm{Re}\ V(k_1 J-1) \quad -\mathrm{Im}\ V(k_1 J-1)$$

$$\mathrm{Re}\ Z(k_1 J+1) \quad -\mathrm{Im}\ Z(k_1 J+1) \quad \mathrm{Re}\ V(k_1 J+1) \quad -\mathrm{Im}\ V(k_1 J+1)$$

$$\ldots\ \mathrm{Re}\ Z(k_1 J+k_w) \quad -\mathrm{Im}\ Z(k_1 J+k_w) \quad \mathrm{Re}\ V(k_1 J+k_w) \quad -\mathrm{Im}\ (k_1 J+k_w)$$

$$\hat{\alpha}_{k_1 y}^{(J)}\ \hat{\beta}_{k_1 y}^{(J)}\ \hat{\alpha}_{k_1 u}^{(J)}\ \hat{\beta}_{k_1 u}^{(J)} \ldots \hat{\alpha}_{k_L y}^{(J)}\ \hat{\beta}_{k_L y}^{(J)}\ \hat{\alpha}_{k_L u}^{(J)}\ \hat{\beta}_{k_L u}^{(J)})'$$

Summarizing, $\hat{\varepsilon}_{w\hat{\theta}}^{(J)}$ may simply be considered as a function of $\hat{\eta}_c^{(J)}$. The computation of the asymptotic expectation and the asymptotic covariance matrix of $\hat{\varepsilon}_{w\hat{\theta}}^{(J)}$ will be based on the asymptotic expectation and asymptotic covariance matrix of $\hat{\eta}_c^{(J)}$. Therefore the latter quantities will first be computed.

According to (1.3.2)

$$Z(k) = \frac{1}{JT} \int_0^{JT} Z(t) \exp(-j2\pi kt/JT)\, dt$$

Hence

$$E(Z(k) = \frac{1}{JT} \int_0^{JT} y(t) \exp(-j2\pi kt/JT)\, dt \qquad (3.6.6)$$

y(t) is periodic with T. So y(t) is also periodic with JT. If y(t) in (3.6.6) is considered periodic with JT, its only harmonics which are not necessarily zero are those corresponding to harmonic numbers which are integer multiples of J. It then follows from (3.6.6) that

$$E\ Z(k) = \begin{cases} \gamma_{iy} & \text{if } k = iJ, \text{ i integer} \\ 0 & \text{otherwise} \end{cases} \qquad (3.6.7)$$

Similarly,

$$E\ V(k) = \begin{cases} \gamma_{iu} & \text{if } k = iJ,\ i \text{ integer} \\ 0 & \text{otherwise} \end{cases} \tag{3.6.8}$$

Furthermore, according to Section 3.1 the expectation of the last 4L elements of $\hat{\eta}_c^{(J)}$ is described by $E\hat{\alpha}^{(J)}_{k_i y} = \alpha_{ky}$, $E\ \hat{\beta}^{(J)}_{k_i y} = \beta_{k_i y}$, $E\ \hat{\alpha}^{(J)}_{k_i u} = \alpha_{k_i u}$ and $E\ \hat{\beta}^{(J)}_{k_i u} = \beta_{k_i u}$. This completes the computation of $E\ \hat{\eta}_c^{(J)}$. Note that the only elements of $E\ \hat{\eta}_c^{(J)}$ which are not necessarily zero are those corresponding to non-zero harmonics of $y(t)$ and $u(t)$. In what follows $E\ \hat{\eta}_c^{(J)} = \eta_c$; note that η_c is also the asymptotic expectation of $\hat{\eta}_c^{(J)}$

The asymptotic covariance matrix of $\hat{\eta}_c^{(J)}$ is defined as

$$\Sigma_{\hat{\eta}_c\hat{\eta}_c} = \frac{1}{JT}\ V_{\hat{\eta}_c\hat{\eta}_c} \tag{3.6.9}$$

where $V_{\hat{\eta}_c\hat{\eta}_c}$ is defined by

$$V_{\hat{\eta}_c\hat{\eta}_c} = \lim_{J\to\infty} E\ \sqrt{JT}(\hat{\eta}_c^{(J)} - E\ \hat{\eta}_c^{(J)})\ \sqrt{JT}(\hat{\eta}_c^{(J)} - E\ \hat{\eta}_c^{(J)})'$$

Since the computation of $\Sigma_{\hat{\eta}_c\hat{\eta}_c}$ is similar to the computation of $\Sigma_{\hat{\eta}\hat{\eta}}$ discussed in Section 3.1 and Section 3.3 it will not be discussed in detail. Here, as an example, the computation of the asymptotic variance of the elements $\sqrt{JT}$ Re $Z(k)$ and $\sqrt{JT}$ Im $Z(k)$ of $\hat{\eta}_c^{(J)}$ will be given. It follows from (3.1.5) that

$$E\ \{\sqrt{JT}(\text{Re } Z(k_1J+\ell) - E \text{ Re } Z(k_1J+\ell))\}^2 =$$

$$= \frac{1}{2}\int_{-JT}^{JT} R_{hh}(t)\ (1 - |t|/JT)\cos 2\pi(k_1J+\ell)t/JT\ dt +$$

$$- \frac{1}{8\pi(k_1J+\ell)} \int_{-JT}^{JT} R_{hh}(t)\sin 2\pi(k_1J+\ell)|t|/JT \, dt \qquad (3.6.10)$$

Correspondingly it can be shown that the variance of $\sqrt{JT}$ Im $Z(k_1J+\ell)$ is obtained from (3.6.10) by replacing the minus sign in front of the second term by a plus sign. Now assume that $R_{hh}(t)$ is absolutely integrable. It has been shown in Section 3.1 that this is a sufficient condition for convergence of the integrals in (3.6.10). Hence (3.6.10) is finite. If in addition it is assumed that $R_{gg}(t)$ is absolutely integrable it can be shown in the same way that the variances of all other elements of $\sqrt{JT}\ \hat{\eta}_c^{(J)}$ are finite. Again assuming that all central moments of $\hat{\eta}_c^{(J)}$ of order higher than two are of order of magnitude lower than $\frac{1}{JT}$, it then follows from Goldberger's result that the asymptotic expectation of any differentiable function of $\hat{\eta}_c^{(J)}$ is the function at $\hat{\eta}_c^{(J)} = \eta_c$, η_c being the asymptotic expectation of $\hat{\eta}_c^{(J)}$. In view of (3.6.7) and (3.6.8) the elements of η_c are either Fourier coefficients of u(t) and y(t) or zero. It then follows from (3.6.2), (3.6.3) and (3.6.4) that η_c satisfies the set of equations

$$\left\{(\hat{P}_w^{(J)}c - \hat{d}_w^{(J)})_{\theta=\hat{\theta}_I^{(J)}}\right\}_{\hat{\eta}_c^{(J)} = \eta_c} = 0 \qquad (3.6.11)$$

So it is concluded from (3.6.5) and (3.6.11) that the asymptotic expectation of $\hat{\varepsilon}_{w\hat{\theta}}^{(J)}$ is zero.

Returning to the computation of $\Sigma_{\hat{\eta}_c\hat{\eta}_c}$ it follows from (3.6.10) that for finite fixed k_w

$$\lim_{J\to\infty} E\,(\sqrt{JT}\ \mathrm{Re}\ Z(k_1J+\ell) - E\ \sqrt{JT}\ \mathrm{Re}\ Z(k_1J+\ell))^2$$

$$= \frac{1}{2} \int_{-JT}^{JT} R_{hh}(t)\ (1 - |t|/JT)\cos 2\pi(k_1J+\ell t/JT\ dt$$

$$(3.6.12)$$

Under the additional assumption that $t\ R_{hh}(t)$ is absolutely integrable it follows from (3.6.12) that

$$\lim_{J\to\infty} E\,(\sqrt{JT}\ \mathrm{Re}\ Z(k_1J+\ell) - E\ \sqrt{JT}\ \mathrm{Re}\ Z(k_1J+\ell))^2$$

$$= \frac{1}{2}\ S_{hh}\ (j2\pi(k_1J+\ell)/JT)$$

Note that this result is very similar to (3.1.10) and can be obtained by substituting $j2\pi(k_1J+\ell)/JT$ for $j2\pi k/T$ in (3.1.10). It can be shown that in the same way all elements of $V_{\hat{n}_c\hat{n}_c}$ can be obtained from the results of Section 3.1 and Section 3.3. The results of these sections are summarized in the expressions (3.3.4). Using these results $V_{\hat{n}_c\hat{n}_c}$ can be written

$$V_{\hat{n}_c\hat{n}_c} = \mathrm{diag}\ (W_{(k_1J-k_w)/J} \cdots W_{(k_1J-1)/J}\ W_{(k_1J+1)/J}$$

$$\cdots W_{(k_1J+k_w)/J}\ W_{k_1} \cdots\cdots W_{k_L}\ )$$

where

$$W_{i/J} =$$

$$\frac{1}{2}\begin{pmatrix} S_{hh}(s_{i/J}) & 0 & \mathrm{Re}\, S_{gh}(s_{i/J}) & -\mathrm{Im}\, S_{gh}(s_{i/J}) \\ 0 & S_{hh}(s_{i/J}) & \mathrm{Im}\, S_{gh}(s_{i/J}) & \mathrm{Re}\, S_{gh}(s_{i/J}) \\ \mathrm{Re}\, S_{gh}(s_{i/J}) & \mathrm{Im}\, S_{gh}(s_{i/J}) & S_{gg}(s_{i/J}) & 0 \\ -\mathrm{Im}\, S_{gh}(s_{i/J}) & \mathrm{Re}\, S_{gh}(s_{i/J}) & 0 & S_{gg}(s_{i/J}) \end{pmatrix} \tag{3.6.13}$$

while according to (3.6.9)

$$\Sigma_{\hat{\eta}_c \hat{\eta}_c} = \frac{1}{JT} V_{\hat{\eta}_c \hat{\eta}_c}$$

This completes the computation of $\Sigma_{\hat{\eta}_c \hat{\eta}_c}$.

Using the expression for $\Sigma_{\hat{\eta}_c \hat{\eta}_c}$ the asymptotic covariance matrix $\Sigma_{\hat{\varepsilon}_{w\hat{\theta}} \hat{\varepsilon}_{w\hat{\theta}}}$ of $\hat{\varepsilon}^{(J)}_{w\hat{\theta}}$ is computed as follows. According to Goldberger's result

$$\Sigma_{\hat{\varepsilon}_{w\hat{\theta}} \hat{\varepsilon}_{w\hat{\theta}}} = \left(\frac{\partial \hat{\varepsilon}^{(J)}_{w\hat{\theta}}}{\partial \hat{\eta}^{(J)}_c} \right)'_{\hat{\eta}^{(J)}_c = \eta_c} \Sigma_{\hat{\eta}_c \hat{\eta}_c} \left(\frac{\partial \hat{\varepsilon}^{(J)}_{w\hat{\theta}}}{\partial \hat{\eta}^{(J)}_c} \right)_{\hat{\eta}^{(J)}_c = \eta_c} \tag{3.6.14}$$

The partial derivatives in this expression are computed as follows. Since at $\hat{\eta}^{(J)}_c = \eta_c$, $\hat{\theta}^{(J)}_I = \theta$ and since at $\hat{\theta}^{(J)}_I = \theta$, $\hat{\varepsilon}^{(J)}_{w\hat{\theta}} = \hat{\varepsilon}^{(J)}_w$

$$\left(\frac{\partial\hat{\varepsilon}^{(J)}_{w\hat{\theta}}}{\partial\hat{\eta}^{(J)}_{c}}\right)'_{\hat{\eta}^{(J)}_{c}=\eta_c} = \left(\frac{\partial\hat{\varepsilon}^{(J)}_{w}}{\partial\hat{\eta}^{(J)}_{c}}\right)'_{\hat{\eta}^{(J)}_{c}=\eta_c} +$$

$$+\left(\frac{\partial\hat{\varepsilon}^{(J)}_{w}}{\partial\theta}\right)'_{\hat{\eta}^{(J)}_{c}=\eta_c} \left(\frac{\partial\hat{\theta}^{(J)}_{I}}{\partial\hat{\eta}^{(J)}_{c}}\right)'_{\hat{\eta}^{(J)}_{c}=\eta_c} \qquad (3.6.15)$$

Since $\hat{\varepsilon}^{(J)}_{w}$ is a function of $\hat{\eta}^{(J)}_{w}$ only and $\hat{\theta}^{(J)}_{I}$ is a function of $\hat{\eta}^{(J)}$ only. (3.6.15) may be written

$$\left(\frac{\partial\hat{\varepsilon}^{(J)}_{w\hat{\theta}}}{\partial\hat{\eta}^{(J)}_{c}}\right)'_{\hat{\eta}^{(J)}_{c}=\eta_c} = \begin{pmatrix} \left(\frac{\partial\hat{\varepsilon}^{(J)}_{w}}{\partial\hat{\eta}^{(J)}_{w}}\right)_{\hat{\eta}^{(J)}_{w}=\eta_w} \\ 0'_1 \end{pmatrix}' +$$

$$+\left(\frac{\partial\hat{\varepsilon}^{(J)}_{w}}{\partial\theta}\right)'_{\hat{\eta}^{(J)}_{w}=\eta_w} \begin{pmatrix} 0'_2 \\ \left(\frac{\partial\hat{\theta}^{(J)}_{I}}{\partial\hat{\eta}^{(J)}}\right)_{\hat{\eta}^{(J)}=\eta} \end{pmatrix}'$$

$$(3.6.16)$$

where $0'_1$ is a $(4L-4) \times (4k_w+2)$ zero matrix, $0'_2$ is a $8k_w \times (N+M+2)$ zero matrix while η_w denotes the asymptotic expectation of $\hat{\eta}^{(J)}_{w}$. Note that the elements of η_w form a subset of the elements of η_c. Now define

$$G'_w = \left(\frac{\partial \hat{\varepsilon}_w^{(J)}}{\partial \hat{\eta}_w^{(J)}} \right)_{\hat{\eta}_w^{(J)} = \eta_w} \tag{3.6.17}$$

and note that G'_w is obtained by substituting $\hat{\eta}_w^{(J)}$ for $\hat{\eta}^{(J)}$ and $\hat{\varepsilon}_w^{(J)}$ for $\hat{\varepsilon}^{(J)}$ respectively in (3.4.7). So analogous to (3.4.15) and (3.4.16) G'_w is described by

$$G'_w =$$

$$\begin{pmatrix} G'_{(k_1J-k_w)/J} & & & & & \\ & \ddots & & & 0 & \\ & & G'_{(k_1J-1)/J} & & & \\ & & & G'_{(k_1J+1)/J} & & \\ & 0 & & & \ddots & \\ & & & & G'_{(k_1J+k_w)/J} & \\ & & & & & G'_{k_1} \end{pmatrix} \tag{3.6.18}$$

where

$$G'_{i/J} \quad \begin{pmatrix} \mathrm{Re}\ A(s_{i/J}) & \mathrm{Im}\ A(s_{i/J}) \\ \mathrm{Im}\ A(s_{i/J}) & -\mathrm{Re}\ A(s_{i/J}) \\ -\mathrm{Re}\ B_\tau(s_{i/J}) & -\mathrm{Im}\ B_\tau(s_{i/J}) \\ -\mathrm{Im}\ B_\tau(s_{i/J}) & \mathrm{Re}\ B_\tau(s_{i/J}) \end{pmatrix} \tag{3.6.19}$$

Furthermore define

$$\Phi'_w = \left(\frac{\partial \hat{\varepsilon}_w^{(J)}}{\partial \theta} \right)_{\hat{\eta}_w^{(J)} = \eta_w} \tag{3.6.20}$$

Note that Φ'_w is obtained by substituting $\hat{\varepsilon}_w^{(J)}$ for $\hat{\varepsilon}^{(J)}$ and $\hat{\eta}_c^{(J)}$ for $\hat{\eta}^{(J)}$ respectively in (3.4.6). So analogous to (3.4.32) and (3.4.33) Φ'_w is described by

$$\Phi'_w = \Big(\Phi'_{(k_1 J - k_w)/J} \cdots \Phi'_{(k_1 J-1)/J}\ \Phi'_{(k_1 J+1)/J} \cdots$$
$$\cdots \Phi'_{(k_1 J + k_w)/J}\ \Phi'_{k_1} \Big)$$

where

$$\Phi_{i/J} =$$

$$\begin{pmatrix} \mathrm{Re}\ s^{o}_{i/J}\ Z(s_{i/J}) \ \dots\ \mathrm{Re}\ s^{N}_{i/J}\ Z(s_{i/J}) & -\mathrm{Re}\ s^{o}_{i/J}\ V_\tau(s_{i/J}) \\ \mathrm{Im}\ s^{o}_{i/J}\ Z(s_{i/J}) \ \dots\ \mathrm{Im}\ s^{N}_{i/J}\ Z(s_{i/J}) & -\mathrm{Im}\ s^{o}_{i/J}\ V_\tau(s_{i/J}) \end{pmatrix}$$

$$\begin{pmatrix} \dots\ -\mathrm{Re}\ s^{M-1}_{i/J}\ V_\tau(s_{i/J}) & -\mathrm{Re} - s_{i/J}\ V_\tau(s_{i/J}) B(s_{i/J}) \\ \dots\ -\mathrm{Im}\ s^{M-1}_{i/J}\ V_\tau(s_{i/J}) & -\mathrm{Im} - s_{i/J}\ V_\tau(s_{i/J}) B(s_{i/J}) \end{pmatrix}_{\hat{\eta}_c^{(J)} = \eta_c}$$

It has been shown above that the only elements of η_c which are not zero are those corresponding to the non-zero harmonics of u(t) and y(t). So the only $\Phi_{i/J}$ which are not null matrices are those corresponding to the frequencies of these harmonics. Finally, according to (3.4.8)

$$\left(\frac{\partial \hat{\theta}_I^{(J)}}{\partial \hat{\eta}^{(J)}}\right)'_{\hat{\eta}^{(J)}=\eta} = -(\Phi'\Phi)^{-1}\Phi' G \qquad (3.6.21)$$

Substituting (3.6.17), (3.6.20) and (3.6.21) in (3.6.16) yields

$$\left(\frac{\partial \hat{\varepsilon}_{w\hat{\theta}}^{(J)}}{\partial \hat{\eta}_c^{(J)}}\right)'_{\hat{\eta}_c^{(J)}=\eta_c} = \begin{pmatrix} G_w' \\ O_1' \end{pmatrix}' + \Phi_w \begin{pmatrix} O_2' \\ -G'\Phi(\Phi'\Phi)^{-1} \end{pmatrix}' \qquad (3.6.22)$$

Hence from (3.6.14) and (3.6.22)

$$\Sigma_{\hat{\varepsilon}_{w\hat{\theta}}\hat{\varepsilon}_{w\hat{\theta}}} = \left\{(G_w \ \ O_1) + \Phi_w(O_2 \ \ -(\Phi'\Phi)^{-1}\Phi' G)\right\}\Sigma_{\hat{\eta}_c\hat{\eta}_c}\left\{(G_w \ \ O_1) + \right.$$
$$\left. + \Phi_w(O_2 \ \ -(\Phi'\Phi)^{-1}\Phi' G)\right\}' = (G_w \ \ O_1)\Sigma_{\hat{\eta}_c\hat{\eta}_c}(G_w \ \ O_1)' +$$
$$+ \Phi_w(O_2 \ \ -(\Phi'\Phi)^{-1}\Phi' G)\Sigma_{\hat{\eta}_c\hat{\eta}_c}(G_w \ \ O_1)' +$$
$$+ (G_w \ \ O_1)\Sigma_{\hat{\eta}_c\hat{\eta}_c}(O_2 \ \ -(\Phi'\Phi)^{-1}\Phi' G)'\Phi_w' +$$
$$+ \Phi_w(O_2 \ \ -(\Phi'\Phi)^{-1}\Phi' G)\Sigma_{\hat{\eta}_c\hat{\eta}_c}(O_2 \ \ -(\Phi'\Phi)^{-1}\Phi' G)'\Phi_w'$$
$$(3.6.23)$$

From the definition of $\hat{\eta}_c^{(J)}$ and $\hat{\eta}_w^{(J)}$ it follows that the $(8k_w + 4) \times 1$ vector $\hat{\eta}_w^{(J)}$ is obtained from the $(8k_w+4L) \times 1$ vector $\hat{\eta}_c^{(J)}$ by leaving out the last $4L-4$ elements of $\hat{\eta}_c^{(J)}$.

Consequently $\Sigma_{\hat{\eta}_w \hat{\eta}_w}$ is the $(8k_w+4) \times (8k_w+4)$ matrix in the upper left hand corner of the $(8k_w+4L) \times (8k_w+4L)$ matrix $\Sigma_{\hat{\eta}_c \hat{\eta}_c}$ and may therefore be written

$$\Sigma_{\hat{\eta}_w \hat{\eta}_w} = \frac{1}{JT} V_{\hat{\eta}_w \hat{\eta}_w}$$

where

$$V_{\hat{\eta}_w \hat{\eta}_w} = \text{diag} \; (W_{(k_1 J - k_w)/J} \cdots W_{(k_1 J-1)/J} \; W_{(k_1 J+1)/J} \cdots W_{(k_1 J+k_w)/J} \; W_{k_1})$$

where $W_{i/J}$ is defined by (3.6.13). From these considerations and from the definition (3.6.17) of G'_w it follows that the first term of (3.6.23) may be written

$$(G_w \;\; O_1)\Sigma_{\hat{\eta}_c \hat{\eta}_c}(G_w \;\; O_1)' = G_w \; \Sigma_{\hat{\eta}_w \hat{\eta}_w} \; G'_w =$$

$$\left(\frac{\partial \hat{\varepsilon}_w^{(J)}}{\partial \hat{\eta}_w^{(J)}}\right)'_{\hat{\eta}_w^{(J)} = \eta_w} \quad \Sigma_{\hat{\eta}_w \hat{\eta}_w} \quad \left(\frac{\partial \hat{\varepsilon}_w^{(J)}}{\partial \hat{\eta}_w^{(J)}}\right)_{\hat{\eta}_w^{(J)} = \eta_w}$$

Hence in view of Goldberger's result

$$(G_w \;\; O_1)\Sigma_{\hat{\eta}_c \hat{\eta}_c} (G_w \;\; O_1)' = \Sigma_{\hat{\varepsilon}_w \hat{\varepsilon}_w} \qquad (3.6.24)$$

It then follows from (3.6.23) and (3.6.24) that the asymptotic covariance matrix of the measured residuals $\Sigma_{\hat{\varepsilon}_{w\hat{\theta}} \hat{\varepsilon}_{w\hat{\theta}}}$ is equal to the sum of the asymptotic covariance matrix of the true residuals $\Sigma_{\hat{\varepsilon}_w \hat{\varepsilon}_w}$ and a number of additional terms. So it is concluded that the squares of the measured residuals, that is the squares of the elements of $\hat{\varepsilon}_{w\hat{\theta}}^{(J)}$, are generally biased estimators of the variances of the corresponding residuals. On the other hand it has been shown above that the only elements of Φ_w which are not necessarily zero are those in the rows corresponding to non-zero harmonics of the test signal. So the only diagonal elements of $\Sigma_{\hat{\varepsilon}_{w\hat{\theta}}\hat{\varepsilon}_{w\hat{\theta}}}$ differing from the corresponding diagonal elements of $\Sigma_{\hat{\varepsilon}_w\hat{\varepsilon}_w}$ are those corresponding to the frequencies of the non-zero harmonics. This is equivalent with the observation that the only asymptotic variances of the measured residuals differing from the asymptotic variances of the true residuals are those corresponding to the frequencies of the non-zero harmonics. So for all other frequencies the square of the measured residual is asymptotically an unbiased estimator of the variance of the corresponding true residual. Therefore, instead of $\hat{\lambda}_w(\omega_{k_1})$ a somewhat modified estimator $\hat{\lambda}_{w'}(\omega_{k_1})$ is proposed, defined by

$$\hat{\lambda}_{w'}(\omega_{k_1}) = \frac{JT}{2(2k_{w'}-k_h)} \sum_{\substack{k=k_1J-k_{w'} \\ k/J \neq \text{ harmonic numbers} \\ \text{of non-zero harmonics}}}^{k_1J+k_{w'}} \mathrm{Re}^2\{A(j2\pi k/JT)Z(k) +$$

$$- B(j2\pi k/JT)\exp(-j2\pi k\tau/JT)V(k)\}_{\theta=\hat{\theta}_I(J)} +$$

$$+ \mathrm{Im}^2\{A(j2\pi k/JT)Z(k) +$$

$$- B(j2\pi k/JT)\exp(-j2\pi k\tau/JT)V(k)\}_{\theta=\hat{\theta}_I(J)}$$

where k_h denotes the number of non-zero harmonics of the test signal in the frequency interval $[(k_1J-k_{w'})2\pi/JT, (k_1J+k_{w'})2\pi/JT]$. Note that the only difference between $\hat{\lambda}_w$, described by (3.6.1), and $\hat{\lambda}_{w'}$ is that in the expression for the latter the terms corresponding to non-zero harmonics of the test signal are not present.

The computation of the asymptotic expectation of $\hat{\lambda}_{w'}(\omega_{k_1})$ is straightforward. Define $\hat{\varepsilon}_{w'}^{(J)}$ as the vector obtained by eliminating from $\hat{\varepsilon}_{w}^{(J)}$ all elements corresponding to frequencies of non-zero harmonics of the test signal. Define $\hat{\varepsilon}_{w',\hat{\theta}}^{(J)}$ and $\hat{\eta}_{w'}^{(J)}$ correspondingly. It then follows from the above considerations that asymptotically

$$E\,[\hat{\lambda}_{w'}(\omega_{k_1})] = \frac{JT}{2(2k_{w'}-k_h)}\operatorname{tr}\,\Sigma_{\hat{\varepsilon}_{w',\hat{\theta}}\hat{\varepsilon}_{w',\hat{\theta}}}$$

$$= \frac{JT}{2(2k_{w'}-k_h)}\operatorname{tr}\,\Sigma_{\hat{\varepsilon}_{w'},\hat{\varepsilon}_{w'}} \qquad (3.6.25)$$

Using Goldberger's result

$$\Sigma_{\hat{\varepsilon}_{w'},\hat{\varepsilon}_{w'}} = \left(\frac{\partial\hat{\varepsilon}_{w'}^{(J)}}{\partial\hat{\eta}_{w'}^{(J)}}\right)'_{\hat{\eta}_{w'}^{(J)}=\eta_{w'}} \Sigma_{\hat{\eta}_{w'},\hat{\eta}_{w'}} \left(\frac{\partial\hat{\varepsilon}_{w'}^{(J)}}{\partial\hat{\eta}_{w'}^{(J)}}\right)_{\hat{\eta}_{w'}^{(J)}=\eta_{w'}}$$

$$= G_{w'}\,\Sigma_{\hat{\eta}_{w'},\hat{\eta}_{w'}}\,G'_{w'} \qquad (3.6.26)$$

where

$$G'_{w'} = \left(\frac{\partial\hat{\varepsilon}_{w'}^{(J)}}{\partial\hat{\eta}_{w'}^{(J)}}\right)_{\hat{\eta}_{w'}^{(J)}=\eta_{w'}}$$

while $\eta_{w'}$ denotes the asymptotic expectation of $\hat{\eta}_{w'}^{(J)}$. Note that the elements of $\hat{\eta}_{w'}^{(J)}$ and $\eta_{w'}$ are a subset of the elements of $\hat{\eta}_{w}^{(J)}$ and η_{w} respectively. Also note that $G_{w'}$ is obtained by eliminating from G_w all $G_{i/J}$ corresponding to frequencies of non-zero harmonics of the test signal. See (3.6.18) and (3.6.19). The computation of the product (3.6.26) follows closely the procedure for computation of

$\Sigma_{\hat{\varepsilon}\hat{\varepsilon}} = G\,\Sigma_{\hat{\eta}\hat{\eta}}\,G'$ discussed in Section 3.4 and will therefore not be discussed here. The result of the matrix multiplication (3.6.26) corresponds to (3.4.26), (3.4.27) and (3.4.28) and is described by

$$\Sigma_{\hat{\varepsilon}_{w',\hat{\theta}}\hat{\varepsilon}_{w',\hat{\theta}}} = \frac{1}{JT}\ \mathrm{diag}\,(\lambda(\omega_{(k_1J-k_w)/J}), \ldots, \lambda(\omega_{(k_1J-1)/J}),\ \lambda(\omega_{(k_1J+1)/J}), \ldots, \lambda(\omega_{(k_1J+k_w)/J})) \tag{3.6.27}$$

where $\omega_{i/J} = 2\pi i/JT$ and

$$\lambda(\omega_{i/J}) = \frac{1}{2}\left[\ |A(j\omega_{i/J})|^2\{S_{h_2h_2}(j\omega_{i/J}) + S_{h_3h_3}(j\omega_{i/J})\} + |B_\tau(j\omega_{i/J})|^2\ S_{g_2g_2}(j\omega_{i/J})\right] \tag{3.6.28}$$

while i/J may not be equal to the harmonic number of a non-zero harmonic of the test signal. It then follows from (3.6.25), (3.6.27) and (3.6.28) that asymptotically

$$E\,[\hat{\lambda}_{w'}(\omega_{k_1})] = \frac{1}{2k_{w'}-k_h} \sum_{\substack{k=k_1J-k_{w'} \\ k/J \neq \text{ harmonic numbers} \\ \text{of non-zero harmonics}}}^{k_1J+k_{w'}} [\ |A(j2\pi k/JT)|^2\{S_{h_2h_2}(j2\pi k/JT) + S_{h_3h_3}(j2\pi k/JT)\} + |B_\tau(j2\pi k/JT)|^2 S_{g_2g_2}(j2\pi k/JT)] \tag{3.6.29}$$

CONCLUSIONS

In this research a procedure has been developed for estimating parameters of linear systems from noise corrupted responses to periodic test signals. The developed least squares estimator is extremely simple from a computational standpoint of view; it only involves Fourier analysis of the input output observations and subsequent solution of a set of linear equations for the unknown parameters. The estimator is consistent under mild conditions. Without additional assumptions the consistency is preserved if the system to be investigated is under closed loop control. Furthermore, the estimator offers the possibility to reduce the amount of input output data substantially by using an input consisting of a few harmonics only. The minimum allowable number of harmonics is determined by the requirement that the number of unknown parameters may not exceed twice the number of available harmonics.

In a number of experiments using computer generated data the variance of the proposed estimator has been compared to the minimum variance bound (Cramér-Rao lower bound) on the variance. In the cases considered the efficiencies are all between fifty and hundred percent. Although this may be satisfactory in most cases a procedure has been developed for reducing the variance in an additional step. Results of numerical experiments indicate that this two-step procedure actually achieves the minimum variance bound.

It has been shown how the minimum variance bound can be manipulated by selection of the test signal.

A numerical procedure has been developed for approximate synthesis of periodic two-level signals having specified power spectra.

In conclusion: the feasibility of the proposed procedure has been proved theoretically. The given numerical examples emphasize its usefulness for practical system analysis. Therefore further investigations of the estimator for increased sample sizes and a wider variety of systems and noises, as well as applications to practical systems seem fully justified.

APPENDIX

Let $\check{\theta}^{(J)} = \check{\theta}(\check{\eta}^{(J)})$ be a vector valued estimator whose elements are functions of the vector valued estimator $\check{\eta}^{(J)}$. The superscript J refers to the size of the sample from which $\check{\eta}^{(J)}$ is computed. Denote the pth element of $\check{\theta}^{(J)}$ and the qth element of $\check{\eta}^{(J)}$ by $\check{\theta}_p^{(J)}$ and $\check{\eta}_q^{(J)}$ respectively. Let $\check{\theta}^{(J)}$ be $I_\theta \times 1$ and let $\check{\eta}^{(J)}$ be $I_\eta \times 1$. Define the asymptotic expectation of $\check{\theta}^{(J)}$ and $\check{\eta}^{(J)}$ as

$$\check{\theta} = \lim_{J\to\infty} E\, \check{\theta}^{(J)} \quad \text{and} \quad \check{\eta} = \lim_{J\to\infty} E\, \check{\eta}^{(J)}$$

respectively and denote the pth element of $\check{\theta}$ by $\check{\theta}_p$ and the qth element of $\check{\eta}$ by $\check{\eta}_q$. Now suppose that the Taylor series expansion of $\check{\theta}_p(\check{\eta}^{(J)})$ about $\check{\eta}^{(J)} = \check{\eta}$ converges to $\check{\theta}_p(\check{\eta}^{(J)})$. Then

$$\check{\theta}_p(\check{\eta}^{(J)}) = \check{\theta}_p(\check{\eta}) + \sum_{q=1}^{I_\eta} \left(\frac{\partial \check{\theta}_p(\check{\eta}^{(J)})}{\partial \eta_q^{(J)}} \right)_{\check{\eta}^{(J)} = \check{\eta}} (\check{\eta}_q^{(J)} - \check{\eta}_q) +$$

$$+ \frac{1}{2} \sum_{q_1=1}^{I_\eta} \sum_{q_2=1}^{I_\eta} \left(\frac{\partial^2 \check{\theta}_p(\check{\eta}^{(J)})}{\partial \check{\eta}_{q_1}^{(J)} \partial \check{\eta}_{q_2}^{(J)}} \right)_{\check{\eta}^{(J)} = \check{\eta}} (\check{\eta}_{q_1}^{(J)} - \check{\eta}_{q_1})(\check{\eta}_{q_2}^{(J)} - \check{\eta}_{q_2}) + \dots$$

(A1)

Hence

$$\check{\theta}_p = \lim_{J\to\infty} E\, \check{\theta}_p(\check{\eta}^{(J)}) = \check{\theta}_p(\check{\eta}) +$$

$$+ \frac{1}{2} \sum_{q_1=1}^{I_\eta} \sum_{q_2=1}^{I_\eta} \left(\frac{\partial^2 \check{\theta}_p(\check{\eta}^{(J)})}{\partial \check{\eta}_{q_1}^{(J)} \partial \check{\eta}_{q_2}^{(J)}} \right)_{\check{\eta}^{(J)} = \check{\eta}} \lim_{J\to\infty} E(\check{\eta}_{q_1}^{(J)} - \check{\eta}_{q_1})(\check{\eta}_{q_2}^{(J)} - \check{\eta}_{q_2}) +$$

$$+ \dots$$

(A2)

Define the asymptotic covariance matrix $\Sigma_{\check{\eta}\check{\eta}}$ of $\check{\eta}^{(J)}$ as

$$\Sigma_{\check{\eta}\check{\eta}} = \frac{V_{\check{\eta}\check{\eta}}}{J}$$

where $V_{\check\eta\check\eta}$ is defined by

$$V_{\check\eta\check\eta} = \lim_{J\to\infty} E\, \sqrt{J}(\check\eta^{(J)} - E\,\check\eta^{(J)})\, \sqrt{J}(\check\eta^{(J)} - E\,\check\eta^{(J)})' \qquad (A3)$$

Note that (A3) may be written

$$V_{\check\eta\check\eta} = \lim_{J\to\infty} E\, \sqrt{J}(\check\eta^{(J)} - \check\eta)\ \sqrt{J}(\check\eta^{(J)} - \check\eta)' \qquad (A4)$$

It is assumed that $V_{\check\eta\check\eta}$ is a matrix of finite constants.
Hence

$$\lim_{J\to\infty} \Sigma_{\check\eta\check\eta} = \frac{V_{\check\eta\check\eta}}{J} = 0 \qquad (A5)$$

Under the additional assumption that the moments of the elements of $(\check\eta^{(J)} - \check\eta)$ of order higher than two also vanish, it follows from (A2) - (A5) that

$$\theta_p = \lim_{J\to\infty} E\, \check\theta_p(\check\eta^{(J)}) = \check\theta_p(\check\eta) \qquad (A6)$$

It then follows from (A1) and (A6) that

$$\sqrt{J}(\check\theta_p^{(J)} - \check\theta_p) = \sum_{q=1}^{I_\eta} \left(\frac{\partial\check\theta_p^{(J)}}{\partial\check\eta_q^{(J)}}\right)_{\check\eta^{(J)}=\check\eta} \sqrt{J}(\check\eta_q^{(J)} - \check\eta_q) +$$

$$+ \frac{1}{2}\sum_{q_1=1}^{I_\eta}\sum_{q_2=1}^{I_\eta} \left(\frac{\partial^2\check\theta_p^{(J)}}{\partial\check\eta_{q_1}^{(J)}\,\partial\check\eta_{q_2}^{(J)}}\right)_{\check\eta^{(J)}=\check\eta} \sqrt{J}(\check\eta_{q_1}^{(J)} - \check\eta_{q_1})(\check\eta_{q_2}^{(J)} - \check\eta_{q_2}) + \ldots$$

Hence

$$E\ \sqrt{J}(\check\theta_{p_1}^{(J)} - \check\theta_{p_1})\ \sqrt{J}(\check\theta_{p_2}^{(J)} - \check\theta_{p_2}) =$$

$$E \sum_{q_1=1}^{I_\eta} \left(\frac{\partial\theta_{p_1}}{\partial\check\eta_{q_1}^{(J)}}\right)_{\check\eta^{(J)}=\eta} \sqrt{J}(\check\eta_{q_1}^{(J)} - \check\eta_{q_1}) \sum_{q_2=1}^{I_\eta} \left(\frac{\partial\theta_{p_2}}{\partial\check\eta_{q_2}^{(J)}}\right)_{\check\eta^{(J)}=\check\eta} \sqrt{J}(\check\eta_{q_2}^{(J)} - \check\eta_{q_2}) +$$

$$+\ J\ E\ (\text{terms of order higher than two}) \qquad (A7)$$

Assuming that asymptotically the moment of the elements of

$(\breve{\eta}^{(J)} - \breve{\eta})$ of order higher than two are of order of magnitude lower than $1/J$ it follows from (A3), (A(4) and (A7) that

$$\lim_{J\to\infty} E\ \sqrt{J}(\breve{\theta}^{(J)}_{p_1} - \breve{\theta}_{p_1})\ \sqrt{J}(\breve{\theta}^{(J)}_{p_2} - \breve{\theta}_{p_2})$$

$$= \sum_{q_1=1}^{I_\eta} \sum_{q_2=1}^{I_\eta} \left(\frac{\partial \breve{\theta}^{(J)}_{p_1}}{\partial \breve{\eta}^{(J)}_{q_1}}\right)_{\breve{\eta}^{(J)} = \breve{\eta}} \left(\frac{\partial \breve{\theta}^{(J)}_{p_2}}{\partial \breve{\eta}^{(J)}_{q_2}}\right)_{\breve{\eta}^{(J)} = \breve{\eta}} v_{q_1 q_2, \breve{\eta}\breve{\eta}} \tag{A8}$$

where $v_{q_1 q_2, \breve{\eta}\breve{\eta}}$ is the $q_1 q_2$ element of $V_{\breve{\eta}\breve{\eta}}$. Now (A8) may be rewritten

$$\lim_{J\to\infty} E\ \sqrt{J}(\theta^{(J)}_{p_1} - \breve{\theta}_{p_1})\ \sqrt{J}(\breve{\theta}^{(J)}_{p_2} - \breve{\theta}_{p_2}) =$$

$$= \left(\frac{\partial \breve{\theta}^{(J)}_{p_1}}{\partial \breve{\eta}^{(J)}}\right)'_{\breve{\eta}^{(J)} = \breve{\eta}} V_{\breve{\eta}\breve{\eta}} \left(\frac{\partial \breve{\theta}^{(J)}_{p_2}}{\partial \breve{\eta}^{(J)}}\right)_{\breve{\eta}^{(J)} = \breve{\eta}} \tag{A9}$$

where $\dfrac{\partial \breve{\theta}^{(J)}_{p_i}}{\partial \breve{\eta}^{(J)}}$ is the $I_\eta \times 1$ vector the qth element of which is defined by $\dfrac{\partial \breve{\theta}^{(J)}_{p_i}}{\partial \breve{\eta}^{(J)}_q}$. It follows from (A9) that

$$\lim_{J\to\infty} E\ \sqrt{J}(\breve{\theta}^{(J)} - \breve{\theta})\ \sqrt{J}(\breve{\theta}^{(J)} - \breve{\theta})' =$$

$$= \left(\frac{\partial \breve{\theta}^{(J)}}{\partial \breve{\eta}^{(J)}}\right)'_{\breve{\eta}^{(J)} = \breve{\eta}} V_{\breve{\eta}\breve{\eta}} \left(\frac{\partial \breve{\theta}^{(J)}}{\partial \breve{\eta}^{(J)}}\right)_{\breve{\eta}^{(J)} = \breve{\eta}} \tag{A10}$$

where $\dfrac{\partial \breve{\theta}^{(J)}}{\partial \breve{\eta}^{(J)}}$ is the $I_\eta \times I_\theta$ matrix the qp element of which is defined as $\frac{\partial \check{\theta}_p^{(J)}}{\partial \check{\eta}_q^{(J)}}$. Now define the asymptotic covariance matrix $\Sigma_{\check{\theta}\check{\theta}}$ of $\check{\theta}^{(J)}$ as

$$\Sigma_{\check{\theta}\check{\theta}} = \frac{V_{\check{\theta}\check{\theta}}}{J} \tag{A11}$$

where $V_{\check{\theta}\check{\theta}}$ is defined by

$$V_{\check{\theta}\check{\theta}} = \lim_{J\to\infty} E\ \sqrt{J}(\check{\theta}^{(J)} - E\ \check{\theta}^{(J)})\ \sqrt{J}(\check{\theta}^{(J)} - E\ \check{\theta}^{(J)})' \tag{A12}$$

Note that (A12) may be written

$$V_{\check{\theta}\check{\theta}} = \lim_{J\to\infty} E\ \sqrt{J}(\check{\theta}^{(J)} - \check{\theta})\sqrt{J}(\check{\theta}^{(J)} - \check{\theta})' \tag{A13}$$

It then follows from (A10) - (A13) that

$$\Sigma_{\check{\theta}\check{\theta}} = \left(\frac{\partial \check{\theta}^{(J)}}{\partial \check{\eta}^{(J)}}\right)'_{\check{\eta}^{(J)} = \check{\eta}} V_{\check{\eta}\check{\eta}} \left(\frac{\partial \check{\theta}^{(J)}}{\partial \check{\eta}^{(J)}}\right)_{\check{\eta}^{(J)} = \check{\eta}} \tag{A14}$$

The derivation of this result closely follows the derivation described by Goldberger (1964), pages 122-125. Therefore in this thesis (A14) is referred to as Goldberger's theorem.

REFERENCES

ÅSTRÖM, K.J. and T. BOHLIN (1965), Numerical identification of linear dynamic systems from normal operating records, IFAC symposium Theory of self-adaptive control systems, Teddington, in: *Theory of self-adaptive control systems*, 96 - 111, New York: Plenum Press, 1966.

ÅSTRÖM, K.J. (1967), On the achievable accuracy in identification problems, IFAC symposium Identification in automatic control systems, Prague, paper 1.8.

ÅSTRÖM, K.J. and P. EYKHOFF (1971), System identification - A survey, *Automatica*, 7, 123 - 162.

BLACKMAN, R.B. and J.W. TUKEY (1958), *The measurement of power spectra*, New York: Dover Publications.

CLARKE, D.W. (1967), Generalized-least-squares estimation of the parameters of a dynamic model, IFAC symposium Identification in automatic control systems, Prague, paper 3.17.

EYKHOFF, P. (1974), *System identification*, New York: John Wiley and Sons.

FEDOROV, V.V. and A. PAZMAN (1968), Design of physical experiments (statistical methods), *Fortschritte der Physik*, 16, 325 - 355.

GARFINKEL, R.S. and G.L. NEMHAUSER (1972), *Integer programming*, New York: John Wiley and Sons.

GENTLEMAN, W.M. and G. SANDE (1966), Fast Fourier transforms - for fun and profit, Fall joint computer conference, *AFIPS proceedings*, 11, 563 - 578, Washington: Spartan.

GOLDBERGER, A.S. (1964), *Econometric theory*, New York: John Wiley and Sons.

GUSTAVSSON, I., L. LJUNG and T. SÖDERSTRÖM (1974), Identification of linear multivariable process dynamics using closed loop experiments, Report 7401, Lund Institute of Technology, Division of Automatic Control.

HOFFMANN DE VISME, G. (1971), *Binary sequences*. London: The English Universities Press.

JENKINS, G.M. and D.G. WATTS (1967), *Spectral analysis*, San Francisco: Holden-Day.

JENSEN, J.R. (1959), Notes on the measurement of dynamic characteristics of linear systems, Part III, Servoteknisk forskningslaboratorium, Danmarks tekniske Højskole.

JOSEPH, P., J. LEWIS and J. TOU (1961), Plant Identification in the presence of disturbances and application to digital adaptive systems, *Transactions AIEE*, 80, Part II, 18 - 24.

KENDALL, M.G. and M.A. STUART (1967), *The advanced theory of statistics*, Volume 2, London: Charles Griffin and Company.

LJUNG, L. (1971), Characterisation of the concept of 'persistently exciting' in the frequency domain, Report 7119, Lund Institute of Technology, Division of Automatic Control.

PLANE, D.R. and C. McMILLAN (1971), *Discrete optimization, integer programming and network analysis for management decisions*, Englewood Cliffs: Prentice-Hall.

VAN DEN BOS, A. (1967), Construction of binary multifrequency test signals, IFAC symposium Identification in automatic control systems, Prague, paper 4.6.

VAN DEN BOS, A. (1970), Estimation of linear system coefficients from noisy responses to binary multifrequency test signals, Second IFAC symposium Identification and process parameter estimation, Prague, paper 7.2.

VAN DEN BOS, A. (1973), Selection of periodic test signals for estimation of linear system dynamics, Third IFAC symposium Identification and system parameter estimation, The Hague, paper TT-3, in: *Identification and system parameter estimation*, Amsterdam: North-Holland 1973.

VAN TREES, H.L. (1968), *Detection, estimation and modulation theory*, Part I, New York: John Wiley and Sons.

WILKS, S.S. (1962), *Mathematical statistics*, New York: John Wiley and Sons.

WONNACOTT, R.J. and T.H. WONNACOTT (1970), *Econometrics*, New York: John Wiley and Sons.