Maximum Likelihood Parameter Identification of Flexible Spacecraft

March 1987

Qi Ping Chu
Maximum Likelihood Parameter Identification of Flexible Spacecraft

Qi Ping Chu
LIST OF CONTENTS

ACKNOWLEDGEMENT

ABSTRACT

CHAPTER 0 LIST OF SYMBOLS AND COORDINATE REFERENCE
   FRAME DEFINITIONS 1
   0.1 List of Symbols 1
   0.2 Coordinate Reference Frame Definitions 11

CHAPTER 1 A GENERAL INTRODUCTION 16
   1.1 The Field of Investigation 16
   1.2 Problem Statement 20
   1.3 Organization of the Thesis 22
   References 23

CHAPTER 2 MATHEMATICAL MODELLING OF FLEXIBLE SPACECRAFT DYNAMICS 27
   2.1 Introduction 27
   2.2 The Kinetic Energy of the Flexible Spacecraft 29
      2.2.1 The kinetic energy of the rigid main body 29
      2.2.2 The kinetic energy of the flexible substructures 35
      2.2.3 The total kinetic energy of the spacecraft 51
   2.3 The Potential Energy of the Flexible Spacecraft 52
      2.3.1 The potential energy of the undeformed spacecraft 52
      2.3.2 The potential energy due to the deformations of the flexible substructures 54
      2.3.3 The total potential energy of the spacecraft 58
   2.4 A General Set of Linear Differential Equations of Motion of Flexible Spacecraft Dynamics 58
   2.5 A Particular Dynamical Model of a Spacecraft with Two Symmetrical Solar Panels 62
      2.5.1 Model idealization 62
      2.5.2 Element and nodal numbering algorithm 64
      2.5.3 Assembled model 74
2.6 Parameters to be Identified in the Dynamical Model

2.6.1 Parameters to be identified in the rigid main body

2.6.2 Parameters to be identified in the flexible substructures

2.6.3 Parameters to be identified in the total dynamical model

References

CHAPTER 3 MAXIMUM LIKELIHOOD ESTIMATION OF FLEXIBLE SPACECRAFT PARAMETERS

3.1 Introduction

3.2 The Mathematical Model to be Used for Maximum Likelihood Parameter Identification

3.2.1 Instumentation systems and output equation of flexible spacecraft

3.2.2 Order reduction of the flexible spacecraft dynamical model

3.2.3 Modelling error characteristics of the reduced order model

3.2.4 Discretization of the system model

3.2.5 Parameters to be identified in the total system model

3.3 Maximum Likelihood Parameter Identification of Linear Systems with Correlated Process and Measurement Noises

3.4 Optimization Methods for ML Parameter Identification

3.5 Identifiability of the Flexible Spacecraft Parameters

3.6 Properties of the Maximum Likelihood Parameter Estimator

3.7 Sensitivity Matrix Calculations for the Reduced Order Model

References

CHAPTER 4 SIMULATION EXPERIMENTS AND RESULTS OF ML PARAMETER IDENTIFICATION OF FLEXIBLE SPACECRAFT

4.1 Introduction
4.2 A Summary of the Computer Programs for ML Estimation of the Parameters in Linear Systems with Correlated Process and Measurement Noises 162

4.3 Evaluation of Computer Program in a Simple Simulation Test Case 167
4.3.1 Test example without correlated process and measurement noises 167
4.3.2 Test example with correlated process and measurement noises 177

4.4 Simulation of Flexible Spacecraft Motions and Measurement Data Generation 183

4.5 The Selection of a Priori Values of the Parameters in the Flexible Spacecraft Model 197

4.6 Simulation Results of the Estimation of the Parameter in the Flexible Spacecraft Model 201
4.6.1 Simulation results of the estimation of the parameters using the model with the white process noises 201
4.6.2 Simulation results of the estimation of the parameters using the model with the colored process noises 213

CHAPTER 5 SUMMARY OF CONCLUSIONS AND SUGGESTIONS FOR FURTHER RESEARCH 226

5.1 Summary of conclusions 226
5.2 Suggestions for Further Research 228
Reference 228

APPENDIX A THE LAGRANGE'S EQUATIONS FOR QUASI-COORDINATES 229

APPENDIX B THE MASS AND STIFFNESS MATRICES OF A RECTANGULAR PLATE ELEMENT 240

APPENDIX C THE FISHER INFORMATION MATRIX OF A DETERMINISTIC SYSTEM 249

SAMENVATTING IN HET NEDERLANDS (SUMMARY IN DUTCH) 259
Chapter 0

List of Symbols and Coordinate Reference Frame Definitions

0.1 List of Symbols

\( A \) = assembled coupling matrix for the rotational degrees of freedom of the rigid main body and the flexible substructure degrees of freedom defined by Eq. (2.62)
\( a \) = width of a plate element
\( A_a \) = augmented system matrix defined by Eq. (3.49)
\( A_c \) = condensed coupling matrix defined by Eqs. (3.8) and (3.183)
\( A_i \) = coupling matrix associated with element \( i \) defined by Eq. (2.48)
\( a_1 \) = matrix defined by Eq. (3.41)
\( A_m \) = coupling matrix associated with the master degrees of freedom defined in Eq. (3.184)
\( A_o \) = original high dimensional system matrix defined in Eq. (2.91)
\( A_r \) = reduced system matrix defined in Eqs. (3.16) and (3.30)
\( A_s \) = coupling matrix associated with the slave degrees of freedom defined in Eq. (3.184)
\( b \) = length of a plate element
\( B_a \) = augmented system input matrix defined by Eq. (3.51)
\( b_1 \) = matrix defined by Eq. (3.42)
\( B_o \) = original high dimensional system input matrix defined in Eq. (2.91)
\( B_r \) = reduced order system input matrix defined in Eqs. (3.16) and (3.30)
\( C \) = shape function matrix in the finite element method
$C_r$  
shape function matrix related to the rotational degrees of freedom

$C_t$  
shape function matrix related to the translational degrees of freedom

$C_{s1}$  
observation matrix defined in Eq. (3.32)

$C_{s2}$  
observation matrix defined in Eq. (3.32)

$C_{s*1}$  
observation matrix defined in Eq. (3.1)

$C_{s*2}$  
observation matrix defined in Eq. (3.1)

$D$  
original high dimensional damping matrix, augmented feed-forward matrix

$D_c$  
original high dimensional vector of the assembled degrees of freedom of flexible structures

condensed damping matrix defined in Eqs. (3.10) and (3.181)

$d_{*k}$  
displacement vector of the $k^{th}$ node in the total displacement vector $d$ of the flexible substructures

$d_{i,k}$  
displacement vector at node $k$ of element $i$

$d_{-i}$  
assembled displacement vector of element $i$

$d_{-m}$  
displacement vector related to the master degrees of freedom

$d_r$  
rotational displacement vector of an element

$d_s$  
displacement vector related to the slave degrees of freedom

$d_t$  
translational displacement vector of an element

$E$  
shaping filter matrix defined in Eq. (3.43)

$E$  
Young's modulus of a material

$E_k$  
kinetic energy of a spacecraft

$E_{ka}$  
kinetic energy of flexible substructures

$E_{ka_i}$  
kinetic energy of element $i$ on a flexible structure
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E_{kb}$</td>
<td>kinetic energy of the rigid main body of a spacecraft</td>
</tr>
<tr>
<td>$E_p$</td>
<td>total potential energy of a spacecraft</td>
</tr>
<tr>
<td>$E_{pa}$</td>
<td>potential energy due to structural deformations</td>
</tr>
<tr>
<td>$E_{pa_i}$</td>
<td>potential energy of element $i$ due to deformations</td>
</tr>
<tr>
<td>$E_{pr}$</td>
<td>potential energy of the undeformed spacecraft</td>
</tr>
<tr>
<td>$f$</td>
<td>unknown disturbance vector defined in Eq. (3.30)</td>
</tr>
<tr>
<td>$F_a$</td>
<td>total force vector acting on the flexible substructures defined in Eq. (2.84)</td>
</tr>
<tr>
<td>$F_b$</td>
<td>external force vector on the rigid main body of a spacecraft</td>
</tr>
<tr>
<td>$F_d$</td>
<td>damping force vector of flexible substructures</td>
</tr>
<tr>
<td>$F_f$</td>
<td>distributed force vector acting on a finite element defined in Eq. (2.69)</td>
</tr>
<tr>
<td>$F_i$</td>
<td>discrete force vector acting on a finite element defined in Eq. (2.74b)</td>
</tr>
<tr>
<td>$E_{x_i}$</td>
<td>displacement function matrix in the finite element method</td>
</tr>
<tr>
<td>$F_o$</td>
<td>original high dimensional system matrix</td>
</tr>
<tr>
<td>$F_r$</td>
<td>reduced order system matrix</td>
</tr>
<tr>
<td>$G_a$</td>
<td>augmented noise input matrix</td>
</tr>
<tr>
<td>$G_i$</td>
<td>first order gradient vector of the likelihood function</td>
</tr>
<tr>
<td>$G_r$</td>
<td>matrix defined in Eq. (3.44)</td>
</tr>
<tr>
<td>$G_{x_l}$</td>
<td>matrix defined in Eq. (3.44)</td>
</tr>
<tr>
<td>$H$</td>
<td>augmented observation matrix defined by Eq. (3.55)</td>
</tr>
<tr>
<td>$h$</td>
<td>assumed input level</td>
</tr>
<tr>
<td>$H_r$</td>
<td>reduced order observation matrix defined by Eq. (3.40)</td>
</tr>
<tr>
<td>$I$</td>
<td>spacecraft inertia matrix</td>
</tr>
<tr>
<td>$I_a$</td>
<td>inertia matrix of flexible substructures</td>
</tr>
<tr>
<td>$I_{ab}$</td>
<td>inertia matrix of the rigid main body of a spacecraft</td>
</tr>
</tbody>
</table>
\( I_i \)  
inertia matrix of element \( i \) defined by Eq. (2.51)

\( K \)  
original high dimensional stiffness matrix, steady state Kalman matrix

\( K(k+1) \)  
nonsteady state Kalman gain matrix at time step \( k+1 \)

\( K_c \)  
condensed stiffness matrix defined in Eqs. (3.11) and (3.180)

\( K_i \)  
stiffness matrix of element \( i \)

\( K_{mm} \)  
stiffness matrix related to the master degrees of freedom

\( K_{ms} \)  
coupling matrix between \( K_{mm} \) and \( K_{ss} \)

\( K_p \)  
one stage prediction gain matrix in the Kalman filter for the system with correlated process and measurement noises

\( K_{ss} \)  
stiffness matrix related to the slave degrees of freedom

\( L(\theta) \)  
negative logarithm of the likelihood function

\( M \)  
original high dimensional mass matrix defined by Eq. (2.60)

\( m \)  
total number of degrees of freedom of flexible substructures, dimension of the Fisher information matrix, total number of parameters to be identified in the system

\( M_c \)  
condensed mass matrix defined by Eqs. (3.9) and (3.179)

\( M_i \)  
mass matrix of element \( i \)

\( m_u \)  
dimension of \( u \)

\( M_{mm} \)  
mass matrix related to the master degrees of freedom

\( M_{ms} \)  
coupling matrix between \( M_{mm} \) and \( M_{ss} \)

\( m_r \)  
number of the master degrees of freedom

\( m_s \)  
number of the slave degrees of freedom
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M_{ss}$</td>
<td>mass matrix related to the slave degrees of freedom</td>
</tr>
<tr>
<td>$m_w$</td>
<td>dimension of $w_1$</td>
</tr>
<tr>
<td>$m_z$</td>
<td>dimension of $z$</td>
</tr>
<tr>
<td>$N$</td>
<td>number of selected elements in the flexible substructures, number of observation samples</td>
</tr>
<tr>
<td>$n$</td>
<td>number of nodes in an element</td>
</tr>
<tr>
<td>$N_n$</td>
<td>total number of nodes on flexible substructures</td>
</tr>
<tr>
<td>$0_1$</td>
<td>$n \times 1$ zero vector</td>
</tr>
<tr>
<td>$0_n$</td>
<td>$n \times n$ zero matrix</td>
</tr>
<tr>
<td>$O_{mn}$</td>
<td>$n \times m$ zero matrix</td>
</tr>
<tr>
<td>$P$</td>
<td>steady state covariance matrix of the state estimates</td>
</tr>
<tr>
<td>$P_{i1}$</td>
<td>assembling matrix between the degrees of freedom of element $i$ and the total degrees of freedom of flexible substructures</td>
</tr>
<tr>
<td>$P(k+1</td>
<td>k)$</td>
</tr>
<tr>
<td>$P(k+1</td>
<td>k+1)$</td>
</tr>
<tr>
<td>$P$</td>
<td>parameter vector to be identified in Chapter 2</td>
</tr>
<tr>
<td>$Q$</td>
<td>coupling matrix between the translational displacements and the rotational displacements of the undeformed spacecraft</td>
</tr>
<tr>
<td>$q$</td>
<td>number of total measured variables</td>
</tr>
<tr>
<td>$Q_{a}$</td>
<td>coupling matrix between the translational and rotational displacements of the undeformed substructures</td>
</tr>
<tr>
<td>$q_a$</td>
<td>number of measured acceleration variables</td>
</tr>
<tr>
<td>$Q_{ab}$</td>
<td>coupling matrix between the translational and rotational displacements of the rigid main body</td>
</tr>
<tr>
<td>$Q_{11}$</td>
<td>coupling submatrix defined in Eq. (2.54)</td>
</tr>
<tr>
<td>$q_r$</td>
<td>number of measured translational displacements of the rigid main body</td>
</tr>
</tbody>
</table>
$q_v$ number of measured velocities of the rigid main body

$q_b$ number of measured attitude angles of the rigid main body

$q_w$ number of measured angular velocities of the rigid main body

$R_{b,o,b}$ vector from the origin of the rigid body fixed reference frame to the origin of the flexible substructure local reference frame, represented in the rigid body fixed reference frame

$R_{b,o,1}, R_{b,o,1}$ vector from the origin of the rigid body fixed reference frame to the origin of the flexible substructure local reference frame, represented in the flexible substructure local reference frame and the associated skew symmetric matrix of elements

$e_{b,p,b}, e_{b,p,b}$ vector from the origin of the rigid body fixed reference frame to a particle in the rigid main body, represented in the rigid body fixed reference frame and the associated skew symmetric matrix of elements

$R_{i,1}$ Fisher information matrix in the $i^{th}$ iteration

$R_{i,j}$ element of the Fisher information matrix at the $i^{th}$ row and the $j^{th}$ column

$e_{1,b}$ vector from the origin of the inertial reference frame to the origin of the rigid body fixed reference frame

$e_{1,p}$ vector from the origin of the inertial reference frame to a particle in the rigid main body of spacecraft

$R_{o,po,b}$ vector from the origin of the flexible substructure local reference frame to the undeformed particle, represented in the rigid body fixed reference frame

$R_{o,po,l}, R_{o,po,l}$ vector from the origin of the flexible substructure local reference frame to the undeformed particle, represented in the flexible substructure local reference frame and the associated skew symmetric matrix of elements
\( R_{p_0,p,b} \) vector from the undeformed particle to the deformed particle, represented in the rigid body fixed reference frame

\( R_{p_0,p,l} \) vector from the undeformed particle to the deformed particle, represented in the flexible substructure local reference frame

\( T \) transformation matrix from the original high dimensional vector of degrees of freedom to the condensed vector of degrees of freedom

\( t \) thickness of a plate element

\( T^* \) transformation matrix between the master degrees of freedom and the slave degrees of freedom

\( T_b \) external torque vector on the rigid main body of a spacecraft

\( T_{b,l} \) transformation matrix from the rigid body fixed reference frame to the flexible substructure local reference frame

\( T_e \) elementary transformation matrix defined in Eq. (3.3)

\( T_{i,o} \) transformation matrix from the inertial reference frame to the orbital reference frame defined by Eq. (0.2)

\( T_{o,b} \) transformation matrix from the orbital reference frame to the rigid body fixed reference frame defined by Eq. (0.3)

\( u \) deterministic input vector

\( U_n \) \( n \times n \) unit matrix

\( v \) measurement noise vector defined by Eq. (3.57)

\( V_i \) matrix defined in Eq. (2.47)

\( V_e \) steady state covariance matrix of the prediction errors (innovations)

\( V_e(k+1|k) \) nonsteady state covariance matrix of the prediction errors (innovations)
\( V_w \) covariance matrix of process noises defined by Eq. (3.66)

\( V_{wv} \) covariance matrix of process and measurement noises defined by Eq. (3.65)

\( V_{v1} \) covariance matrix of measurement noises defined in Eq. (3.59)

\( y_1 \) measurement noise vector defined in Eq. (3.32)

\( \hat{W} \) coupling matrix between the translational degrees of freedom of the rigid main body and the flexible substructure degrees of freedom defined by Eq. (2.63)

\( \hat{W} \) estimated process noise vector

\( \hat{W}_i \) \( i \)th submatrix of matrix \( \hat{W} \) defined by Eq. (2.46)

\( x(k) \) state variable vector at the \( k \)th time step

\( \hat{x}(k+1|k) \) one stage prediction of states

\( \hat{x}(k+1|k+1) \) state estimate vector

\( \hat{y}(k+1|k) \) prediction error (innovation) vector

\( y(k) \) measurement variable vector at the \( k \)th time step

\( \hat{y}(k+1|k) \) one stage prediction output vector

\( z \) state variable vector of a shaping filter defined in Eq. (3.43)

\( \alpha \) damping coefficient related to the mass matrix defined in Eq. (2.86)

\( \alpha(i) \) factor generated by a line minimization search in the Gauss-Newton optimization procedure at the \( i \)th iteration
\( \delta \)
damping coefficient related to the stiffness matrix defined in Eq. (2.86)

\( \Gamma_{uw} \)
deterministic input distribution matrix defined by Eq. (3.68)

\( \Gamma_{w} \)
optimal input distribution matrix defined by Eq. (3.69)

\( \hat{\delta} \)
offset angle of flexible appendages in Chapter 2

\( \zeta_i \)
damping ratio of the \( i^{th} \) mode defined in Eq. (2.88)

\( \theta \)
pitch angle of spacecraft rigid main body

\( \hat{\theta} \)
parameter to be identified in Chapter 3, attitude angle vector defined in Eqs. (2.82) to (2.84)

\( \hat{\phi} \)
maximum likelihood estimates of parameters

\( \lambda_j \)
\( j^{th} \) eigenvalue of Fisher information matrix

\( \nu \)
Poisson ratio of a material

\( \rho \)
mass density of the material of flexible substructures

\( \sigma^2_{v,1,j} \)
element of the variance matrix \( \Sigma_v \) at the \( i^{th} \) row and \( j^{th} \) column

\( \sigma^2_{w,1,j} \)
element of the variance matrix \( \Sigma_w \) at the \( i^{th} \) row and \( j^{th} \) column

\( \Phi \)
transition matix of the system matrix defined by Eq. (3.67)

\( \phi \)
roll angle of the spacecraft rigid main body

\( \Psi \)
matrix defined by Eq. (3.140)

\( \psi \)
yaw angle of the spacecraft rigid main body

\( \omega_{b,i}, \omega_{b,i} \)
angular velocity vector of the spacecraft rigid body

\( \omega \)
fixed reference frame relative to the inertial reference frame and the associated skew symmetric matrix of elements represented in the rigid body fixed reference frame.

\( \omega_o \)
anglular velocity of the spacecraft orbit

\( \omega_o \)
skew symmetric matrix of \( \omega_o \) defined in Eqs. (2.82) to (2.84)
\[ \int_{E_i}^{} \, dm \quad \text{integration with respect to mass within the } i^{th} \text{ element of a flexible substructure} \]
\[ \int_{b}^{} \, dm \quad \text{integration with respect to mass within the rigid main body of the spacecraft} \]
\[ \int_{E_i}^{} \, dv \quad \text{integration with respect to volume within the } i^{th} \text{ element of a flexible substructure} \]
\[ \sum_{i=1}^{N} a_i \quad \text{summation of } a_i \text{ sine } i = 1 \text{ till } i = N \]
\[ \prod_{i=1}^{N} a_i \quad \text{multiplication of } a_i \text{ sine } i = 1 \text{ till } i = N \]
0.2 Coordinate Reference Frame Definitions

The reference frames to be used in this study are as follows.

1. The rigid main body fixed reference frame $F_b$.
We define first a rigid main body fixed reference frame system, which is represented by the unit vector triad $(\hat{i}_b, \hat{j}_b, \hat{k}_b)$ with origin $O_b$. This body axis system will be assumed to be fixed to the rigid main body of the flexible spacecraft. The origin $O_b$ is fixed in the rigid main body.

2. The orbital reference frame $F_o$.
The orbital reference frame is defined by the unit vector triad $(\hat{i}_o, \hat{j}_o, \hat{k}_o)$. The origin of the orbital reference frame $O_o$ is located on the nominal spacecraft orbit, $\hat{i}_o$ points to the orbital direction, $\hat{k}_o$ points to the center of the earth and $\hat{j}_o$ completes the orthogonal system (normal to the orbit plane).

3. The inertial reference frame $F_i$.
The fundamental coordinate system, to which all motions must be referred, is the inertial reference frame. In its most general sense, it is a coordinate system fixed with respect to the stars. However, practical situations dictate only, that the inertial frame be a reference coordinate set which guarantees the required accuracy over the time interval of interest. For the problem considered in this study, it is sufficient to select a coordinate frame with origin $O_i$ at the origin of the orbital reference frame $O_o$ and one axis directed along a fixed celestial direction, such as the first point of Aries. one other axis would be normal to the orbital plane and the third axis would complete the orthogonal set (in the orbit plane). The inertial reference frame can be represented by the unit vector triad $(\hat{i}_i, \hat{j}_i, \hat{k}_i)$ with origin $O_i$ (see Fig. 0.1).

In particular, when the discussed flexible spacecraft has symmetrical structure, the origin of the inertial reference frame can be considered as the
origin of the body fixed reference frame, i.e., the translational displacement of the spacecraft from the origin of the orbital reference frame $O_o$ to the origin of the body fixed reference frame $O_b$ remains zero (see Ref. 7 and Chapter 2 in this study). This inertial reference frame will be considered for a special symmetric flexible satellite in this study as well.

The transformations from the inertial reference frame $I_i$ to the orbital reference frame $O_o$ and from the orbital reference frame $O_o$ to the rigid body fixed reference frame $I_b$ are now required. When defining the orientation of a body with respect to a reference frame, a series of pure rotations is used, and this results in an orthogonal transformation. The associated rotations are called Euler angles and they uniquely determine the orientation of the body. For the transformation from the orbital reference frame $O_o$ to the rigid body fixed reference frame $I_b$, the Euler angles are defined as follows:

1) rotation about the $\vec{k}_o$ axis by the yaw angle $\psi$, positive in the clockwise direction,

2) rotation about the intermediate $\vec{j}$ axis by the pitch angle $\theta$, positive in the clockwise direction, and

3) rotation about the final $\vec{i}$ axis by the roll angle $\phi$, positive in the clockwise direction.

The inertial reference frame $I_i$, the orbit reference frame $O_o$ and the body fixed reference frame $I_b$ are presented in Fig. 0.1.

The transformations can be obtained from Fig. 0.1 as:

\[
\begin{bmatrix}
\vec{i}_b \\
\vec{j}_b \\
\vec{k}_b
\end{bmatrix} = T_{o,b} \cdot T_{i,o} \cdot 
\begin{bmatrix}
\vec{i}_i \\
\vec{j}_i \\
\vec{k}_i
\end{bmatrix}, \quad (0.1)
\]
where $T_{i,0}$ is the transformation matrix between the inertial reference frame $F_i$ and the orbital reference frame $F_0$, $T_{o,b}$ is the transformation matrix between the orbital reference frame $F_o$ and the rigid body fixed reference frame $F_b$, and $\hat{i}_b \hat{j}_b \hat{k}_b$ and $\hat{i}_i \hat{j}_i \hat{k}_i$ are the components of a unit vector in the body fixed reference frame and the inertial reference frame, respectively.

The matrix $T_{i,0}$ is given by:

$$
T_{i,0} = \begin{bmatrix}
\cos(\omega_0 t + \theta_0) & 0 & \sin(\omega_0 t + \theta_0) \\
0 & 1 & 0 \\
-\sin(\omega_0 t + \theta_0) & 0 & \cos(\omega_0 t + \theta_0)
\end{bmatrix}
$$  (0.2)

in which $\omega_0$ is the orbital angular velocity and $\theta_0$ is a constant angle.

The matrix $T_{o,b}$ is given by:

$$
T_{o,b} = \begin{bmatrix}
\cos\theta \cos\phi & \cos\theta \sin\phi & -\sin\theta \\
\sin\phi \sin\theta \cos\psi + \cos\phi \sin\psi & \sin\phi \sin\theta \sin\psi + \cos\phi \cos\psi & \sin\phi \cos\theta \\
\cos\phi \sin\theta \sin\psi - \sin\phi \cos\psi & \cos\phi \sin\theta \sin\psi - \sin\phi \cos\psi & \cos\phi \cos\theta
\end{bmatrix}
$$  (0.3)

If the Euler angles are very small, the transformation matrix $T_{o,b}$ can be simplified as:

$$
T_{o,b} = \begin{bmatrix}
1 & \psi & -\theta \\
-\psi & 1 & \phi \\
\theta & -\phi & 1
\end{bmatrix}
$$  (0.4)

4. The flexible substructure local reference frame $F_1$.

To develop the equations of motion of the flexible substructures of the spacecraft, we define certain flexible substructure local reference frames as follows.

For any flexible substructure, the origin $O_1$ is fixed at some position in the substructure, and the unit vectors $\hat{i}_1, \hat{j}_1$ and $\hat{k}_1$ point to suitable directions (see Fig. 0.3).
The transformation from the body fixed reference frame $F_b$ to the flexible substructure local reference frame $F_1$ is:

$$
\begin{bmatrix}
\hat{i}_1 \\
\hat{j}_1 \\
\hat{k}_1
\end{bmatrix} = T_{b,1} \cdot 
\begin{bmatrix}
\hat{i}_b \\
\hat{j}_b \\
\hat{k}_b
\end{bmatrix},
$$

where $\hat{i}_1$, $\hat{j}_1$ and $\hat{k}_1$ are the components of a unit vector represented in $F_1$ and $T_{b,1}$ denotes a transformation matrix between $F_1$ and $F_b$.

It should be noted here that in this study the finite element method (see Chapter 2) will be applied to analyse the dynamical motion of the flexible substructures. This means that flexible substructure local reference frame, in fact, will be located in each element to be discussed.

Fig. 0.1 The inertial reference frame $F_1^*$, the orbital reference frame $F_0^*$ and the body fixed reference frame $F_b$ definitions.
Fig. 0.2 Transformation from the inertial reference frame $F_1$ to the body fixed reference frame $F_b$.

Fig. 0.3 The flexible substructure local reference frame $F_1$ definition.
CHAPTER 1

A GENERAL INTRODUCTION

1.1 The Field of Investigation

In the early days of space exploration, when spacecraft were small, mechanically simple and relatively compact, they were considered as rigid bodies for the purpose of predicting their dynamical behaviour in space. This practice, however, was soon found to be inadequate in many cases. Explorer I, as the classical example, did not persist in the intended state of spin about its axis of symmetry, but soon tumbled over. The explanation for this anomalous behaviour lay in the flexibility of the small wire turnstile antennas protruding from the cylindrical housing of the vehicle, see [3]. Had the satellite been a truly rigid body, it would have maintained the spinning motion imparted to it during launch. Because of energy dissipation induced by the motion of the flexible antennas, the spinning motion was unstable and the satellite ended up rotating about its symmetric axis, i.e. the axis which is normal to the longitudinal axis.

Great strides in the understanding of the dynamics of flexible spacecraft have been made since 1958, when Explorer I gave an early warning of the dangers of treating space vehicles as rigid bodies. However, despite the best efforts of engineers, numerous surprise performances have been given. In addition to Explorer I, the following satellites, are examples.

- Aloutte I (1962), where rapid spin decay due to solar torque on the thermally deformed vehicle was experienced,
- 1963-22A a gravity gradient stabilised spacecraft was subjected to excessive librations, due to boom bending arising from solar heating,
- Explorer XX (1964), which behaved in the same way as Aloutte I,
- OGO III (1966), controlled by reaction wheels, which developed excessive attitude oscillations due to control system interaction with flexible booms,
- OV1-10 (1966), which behaved similarly to 1963-22A, and
- TACSat I (1969), a dual spin stabilized vehicle, which demonstrated an unexpected limit cycle due to energy dissipation in the bearing assembly.
The examples of 'flexible' spacecraft quoted so far have been extreme in their behaviour. Nevertheless, they have indicated the meticulous attention which must be brought to bear on any new spacecraft design in terms of dynamic analysis at the design stage, if disasters are to be avoided. The current trend is toward increasingly flexible spacecraft, as is indicated by the following examples.

The Canadian Communication Technology Satellite (CTS) carries two solar panels 1.2m x 7.3m, to generate 1.2kw electric power. The Radio Astronomy Explorer (RAE) Satellite used four 230m antennas for detecting low frequency signals. Among the European spacecraft exhibiting flexibility, GEOS, ISEE-B and OTS, can be mentioned. From these examples it may be concluded that modern spacecraft consist of structural subsystems, some essentially rigid, such as the rigid main body of the satellite, and some flexible, such as solar arrays, antennas, etc. It also can be expected that in the future larger and larger spacecraft will be placed in orbit. Due to mass limitation, these space vehicles will be extremely flexible. Examples are the NASA space station and the European space station (COLUMBUS).

When spacecraft are considered as flexible space structures, the difficulties to analyse the system will be increased, because of the complexity of the vehicle dynamics. The main problem areas are:

1. Dynamical modelling of the given flexible space structure.
2. Determining the system uncertainties or parameters.
3. Designing the control system.

The problems of flexible spacecraft will be of major concern in space attitude control in many forthcoming missions. Therefore, a large amount of work related to the areas of modelling and control of flexible spacecraft is being done by many space researchers. The comprehensive survey publication by Modi, see [3], alone mentions over 200 publications associated with modelling and control of flexible structures in space. However, the identification of parameters of models of flexible spacecraft is addressed relatively rarely. There appears to be only a very small number of references in which this particular identification problem is addressed in some detail.

The parameter identification of a flexible spacecraft is a subject as important as the parameter identification of a aircraft, which has been studied since 1960's, see the works of Gerlach [1] and [2] (1964, 1970), Mulder [17]
and [18] (1973, 1979), Mehra [20] (1973), etc. The reason for the importance of parameter identification for a flexible spacecraft is, that the parameters of a flexible spacecraft model might be changeable in space, since:

1. the mass of the spacecraft may be dissipated due to the operation of the control system,
2. the structure of the spacecraft may be deployed or reconfigured after launch,
3. the shape of the spacecraft, e.g. long beam antennas or other flexible appendages, may change due to the surrounding conditions of the spacecraft, such as imposed by solar heating, etc.

All the effects, as mentioned above, lead to changes of the parameters of the flexible spacecraft system model. To design a control system of the spacecraft, the system model to be used should be as accurate as possible, since the performance of the control system depends on the accuracy of the model. This means that the parameters in the system model should also be accurate.

Many papers on general system identification have been presented. The works of Gupta (1981), Hendricks, Rajaram, Kamat and Junkins (1984), Taylor (1985), Chen, Kuo and Garba (1983), Heylen and van Honacker (1983) and Berman (1979, 1983) are relevant to the study presented in this report.

Gupta [4] gives an introduction in which the system identification of flexible space structures is mentioned. Hendricks [5] finds mass, stiffness and damping matrices, from the measurements of acceleration, velocity and displacement at all degrees of freedom, using the least squares method. Taylor, see [6], uses the deterministic maximum likelihood estimation method to find the unknown parameters and suggests using partial differential equations to construct the dynamical model, in order to reduce the number of unknown parameters. Chen, see [7], [8], discusses an iterative approach, based on calculating the Jacobian matrix, which measures the sensitivities of the system parameters. Berman, see [9], [10], finds mass and stiffness matrices 'nearest' to the a priori model matrices, which satisfy certain orthogonality constraints with respect to measured frequencies and mode shapes. Heylen, see [12], follows the approach of Berman to obtain a first correction to the system matrices. The correction is augmented by a second modification, using eigenvalue sensitivities.

All the papers mentioned above, seek best estimates for the elements of the system mass, damping and stiffness matrices or the canonical forms (modal
coordinates) for these matrices. Unfortunately, when the original flexible spacecraft of continua is discretized by using the finite element method, the order of the dynamical models will usually be very high (higher than 200th is common), except for very simple structures with only few elements. Therefore, large dimensional matrices of mass, damping and stiffness will result in the dynamical models of flexible spacecraft. If all elements of these matrices have to be estimated, the number of these parameters will become unmanageable, see [4], [5], and [6], even when the canonical forms (modal coordinates) are used to reduce the number, see [6]. Taylor, see [6], suggests that the flexible spacecraft dynamics should be described by partial differential equations, to greatly reduce the number of unknown parameters. But, the partial differential equations may be converted to ordinary differential equations and the distributed parameters in the partial differential equation will be discretized as a large number of constant parameters during the conversion, see [26]. Moreover, most of the algorithms in parameter identification, state estimation and system optimal control design are based on the system model in ordinary differential equation form, see e.g. [4], [5], [6], [7], [8], [9], [10], [11], [12], [24], [25], [26], [28] and [29]).

On the other hand, from the control system design point of view, very high order models will always be very difficult to use to design the control systems, due to limitations in computer word length, memory size, speed and numerical accuracy. In other words, the order of the originally high order system must be reduced for control system design.

If the order of the system is reduced, certain modelling errors will creep up in the reduced order model. The parameter estimation algorithms must handle the system with large modelling errors. The references, mentioned above, are all working with the system model of high order and assume that there are no modelling errors in the system model. The models to be used for parameter identification are actually deterministic models of high order. However, these papers give a possibility to identify the parameters for high order system models. Once the parameters are obtained by using these algorithms for a high order system model, that model can be used to reduce the order of the system for system control design.

In a system model of reduced order, modelling errors are unavoidable. The identification algorithms, as used for deterministic systems addressed in
the literature, may not treat this particular problem since they only work well in systems without modelling errors or with very small modelling errors, see [4], [13] and [18].

1.2 Problem Statement

From Section 1.1, it may be concluded that the area of parameter identification of flexible spacecraft is an important subject of research. According to the discussions in Section 1.1, the main problems addressed in the literature, possibly to be solved in this area are as follows:

1. The dynamical model of a flexible spacecraft should be a lower order system, not only for the parameter estimation but also for the state estimation and system control design.
2. The number of parameters to be estimated should be reasonably low, such that it can be managed by the parameter estimation algorithm.
3. The parameter estimation algorithms must be able to handle the system with modelling errors resulting from order reduction. In other words the algorithms must be insensitive to modelling errors.
4. The characteristics of the obtained model should be a sufficiently accurate reflection of the characteristics of the real flexible spacecraft. The criterion to measure the correspondence should not only be based on the estimated model parameters but also on the system state estimates.

The above mentioned problems may be characterized as problems occurring mainly due to computational limitations in the application of system identification.

One possible classification of identification methods is based on the type of implementation, see [13], viz.:

1. explicit methods, and
2. parameter adjustment methods.

This classification is useful from the computation point of view, especially with respect to the costs. In general, explicit identification methods are less accurate than parameter adjustment method, but are computationally more
efficient and some times are indispensable to obtain initial parameter estimates, see [14].

This study will focus on the maximum likelihood estimation method, one of the advanced parameter adjustment methods, to estimate the parameters of complex systems.

The method of maximum likelihood was developed by Fisher (1912, 1921). Although the basic ideas date back to Gauss (1809), the method has been further expanded by many researchers over the last 20 years, see e.g. [13], [14], [15], [16], [17], [18], [19], [20], [21], and [22].

The reason of using the maximum likelihood method to estimate the flexible spacecraft parameter may be summarized as follows.

1. It is a very general method for parameter estimation (Aström [15]), since the parameters estimated from this method include structural parameters of flexible spacecraft, initial conditions of system state variables as well as the noise statistics (means, variances and covariances of measurements and process noises). The state estimation of a flexible spacecraft can also be obtained from a Kalman filter in the maximum likelihood approach.

2. The maximum likelihood method provides a lower bound on the variances of the estimated parameters and provides the models of the measurement and process noise disturbances for the means, variances and covariances.

3. Provided that there are no modeling errors or that the modeling errors are correctly modeled, the parameter estimates are consistent and asymptotically unbiased, efficient and normal (Eykhoff [13]).

The extended Kalman filter may also be used to estimate the parameters in the system model, see [4]. However, the properties of the process noise and measurement noise (means, variances and covariances) can not be obtained from the extended Kalman filter.

However, the standard maximum likelihood estimation algorithm, as discussed in the mentioned literature, can not be directly employed in the application to flexible spacecraft, since the measurement instruments or transducers in the flexible spacecraft may consist of accelerometers which detect the specific forces on the flexible appendages rather than the pure kinematic accelerations. The difference between the measured specific force and the required acceleration is the component of gravity along the sensitive axis of the transducer. Assuming the attitude of the transducer relative to the local vertical to vary only very slowly, fluctuations in the specific force
will be very nearly identical to the fluctuations in the kinematic accelerations. Therefore, in the following it will be assumed that the accelerometers measure kinematic accelerations rather than specific forces. When the specific forces on a flexible spacecraft are used as part of the measurements, the process and measurement noises will be correlated, as will be shown in Chapter 3.

Including this particular problem in the flexible spacecraft model, in this report the following work will be done.

1. The finite element method is applied to generate a mathematical model of a flexible spacecraft. Relatively few parameters govern the behaviour of the individual elements in this model. A reduction in the complexity of the calculations is obtained by carefully reducing the order of the model. Finally the parameters to be estimated are obtained by using the maximum likelihood estimation method.

2. Different types of modelling errors are discussed in some detail. They are due to the order reduction and the discretization. Some applicable modelling error models are given.

3. As mentioned above, a suitable order reduction algorithm is used, to reduce the order of the original high dimensional dynamical model of the flexible spacecraft. The result is a reduced order model applicable to parameter identification.

4. Some new algorithms of maximum likelihood estimation are developed, to solve the problems associated with correlated process and measurement noises, based on different considerations.

A limitation of this study is, that only linear system models are considered. This means that only small position and attitude errors of the spacecraft, and small deformations of the flexible substructures of the spacecraft are discussed.

1.3 Organization of the Thesis

The kernel of the thesis consists of the Chapters 2 through 4. In Chapter 2, a mathematical model of general flexible spacecraft dynamics will be obtained, using finite element analysis. To discuss the model in detail, a
particular satellite with two large flexible solar panels will be considered. Large numbers of parameters will occur. But they all depend relatively few key parameters in the dynamical model. As a consequence the number of the parameters to be estimated in the model given in this Chapter remains manageable.

In Chapter 3, the order of the model discussed in Chapter 2 will be reduced by a suitable algorithm. Modelling errors from the order reduction will be investigated. In this chapter, three different algorithms, associated parameter estimation in a system with correlated process and measurement noises, will be obtained based on different considerations.

In Chapter 4, the description of computer programs of the three algorithms will be given. They are followed by some test calculations. Simulated measurement data of the particular flexible spacecraft discussed in Chapter 2 will be generated for parameter estimation. Two different models of modelling errors, caused by the discretization and order reduction, will be applied to the reduced system model and results of parameter estimation will be given.

In addition, some auxiliary results obtained from the main text of this report are given in appendices.

References


CHAPTER 2

MATHEMATICAL MODELLING OF FLEXIBLE SPACECRAFT DYNAMICS

2.1 Introduction

As mentioned in Chapter 1, this study focuses on the parameter identification of flexible spacecraft. The parameters to be identified occur in a mathematical model of the spacecraft. To allow identification, a suitable form of the mathematical model must be developed. The development of such a model is the subject of the present chapter.

According to [1], mathematical models of flexible spacecraft can take either of two different forms:

1) partial differential equations with distributed model parameters (see, e.g., [2], [3], [4] and [5]), and

2) ordinary differential equations (see, e.g., [6] and [7]).

In this study, the second form will be applied, since most algorithms in parameter identification, state estimation and system control design are based on system models in terms of ordinary differential equations (see, e.g., [15], [16], [17], [20], [21], [22], [23] and [24]) and partial differential equations may be converted to ordinary differential equations as well (see, e.g., [19]).

One possibility to generate the mathematical model of a flexible spacecraft in terms of ordinary differential equations is based on the analysis of flexible space structures using the 'finite element method'. Finite element techniques for structural analysis are widely used and accepted. The basic theory, developed and augmented by many researchers over the last twenty years, is available in a number of comprehensive text books, such as [8], [9] and [10].

In the finite element method, a modified structure system consisting of discrete (finite) elements is substituted for the actual continuum. There are several methods to obtain the equations of motion of flexible spacecraft dynamics in finite element form. For example, Likins [6] used Newton's method to obtain both rigid body and flexible appendage equations of motion. Nguyen and Hughes [7] applied the 'angular momentum principle' to obtain the rigid body motion and used Lagrange's equation to establish expressions for
the flexible appendage motion. The equations of motion of flexible spacecraft in this chapter will be more general as compared to the work of Nguyen and Hughes [7], in that the translational displacements of the rigid main body of the spacecraft will be included. Furthermore, the model treated in this chapter can be applied for the spacecraft with any kind of flexible substructures which are not necessarily to be symmetrical structures since the rigid body fixed reference frame is selected not to be the mass center of the rigid main body of the spacecraft, see Section 0.2.

This chapter is organized as follows.

To employ Lagrange's formulation of the equations of motion, expressions for the kinetic energy and potential energy must be obtained. They will be given in Sections 2.2 and 2.3. A general linear dynamical model of a flexible spacecraft will be obtained in section 2.4. The model as derived in section 2.4, is generic in the sense that it is valid not only for two-dimensional flexible structures but also for three-dimensional flexible structures. However, for a spacecraft with two-dimensional-symmetrical panels, the model can be simplified. In this study a satellite with two symmetrical solar panels will also be examined. The model for this particular spacecraft will be discussed in Section 2.5 in some detail.

The selection of the parameters to be identified in flexible spacecraft dynamical models is the final purpose of this chapter. It will be shown that, using the procedure described in this chapter, the number of parameters to be identified can be reduced to a reasonable value. This result is different from other recently available studies in which the authors try to estimate all the elements in the so-called modal matrices, for example, [15], [16] and [17]. The latter studies are expected to lead to many difficulties because of the unmanageably large number of unknown model parameters, even if canonical forms are used [17]. In fact, if the finite element method is employed as the method to obtain the dynamical model of a flexible spacecraft, a large number of the 'unknown parameters' mentioned in [15], [16] and [17] can be obtained directly from the finite element analysis. It will be shown in Chapter 4, that due to this reduction of the number of unknown parameters, the difficulties with respect to parameter identifiability as addressed in [15], [16] and [17] may be circumvented.
2.2 The Kinetic Energy of the Spacecraft

The kinetic energy of a flexible spacecraft consists of the kinetic energy of the rigid main body and the kinetic energy of the flexible substructures. In this section these two parts of the kinetic energy separately will be discussed separately.

2.2.1 The kinetic energy of the rigid main body

Considering a particle $p$ with mass $dm$ in the rigid main body, the kinetic energy of the particle relative to the assumed inertial reference frame, see Section 0.2, can be written as:

$$dE_{\text{kb}} = \frac{1}{2} \ddot{r}_{1,p} \cdot \dot{r}_{1,p} \ dm,$$  \hspace{1cm} (2.1)

where $dE_{\text{kb}}$ is the kinetic energy of the particle, $r_{1,p}$ is a vector from the origin $O_1$ in the inertial reference frame $F_1$ to the discussed particle $p$ (see Fig. 2.1) and the `\cdot` means differentiation with respect to time. From Fig. 2.1, the vector $r_{1,p}$ can also be decomposed as:

$$r_{1,p} = r_{1,b} + r_{b,p},$$  \hspace{1cm} (2.2)

where $r_{1,b}$ is a vector from the origin $O_1$ in the inertial reference frame $F_1$ to the origin $O_b$ in the body fixed reference frame $F_b$ and $r_{b,p}$ is a vector from $O_b$ to the discussed particle $p$ (see Fig. 2.1).

Since $F_b$ is moving relative to $F_1$, the differentiation of $r_{b,p}$ with respect to time in $F_1$ can be represented as:

$$\dot{r}_{b,p} = \dot{r}_{b,p,b} + \omega_{b,i} \times r_{b,p,b},$$  \hspace{1cm} (2.3)
where $\dot{r}_{b,p,b}$ means differentiation of $\dot{r}_{b,p}$ with respect to time in the body fixed reference frame $F_b$, $\dot{r}_{b,p,b}$ is $\dot{r}_{b,p}$ represented in $F_b$ and $\omega_{b,i}$ is the angular velocity vector of $F_b$ with respect to $F_i$.

The first term of (2.3) represents the velocity vector of the particle $p$ with respect to the body fixed reference frame $F_b$.

From Fig. 0.2, the angular velocity vector can be written in $F_b$ as:

$$\omega_{b,i} = \omega_x \dot{i}_b + \omega_y \dot{j}_b + \omega_z \dot{k}_b ;$$

(2.4)

in which $\omega_x$, $\omega_y$ and $\omega_z$ can be obtained from following kinematics:

$$\begin{bmatrix}
\omega_x \\
\omega_y \\
\omega_z
\end{bmatrix} = \begin{bmatrix}
1 & 0 & -\sin\theta \\
0 & \cos\phi & \sin\phi\cos\theta \\
0 & -\sin\phi & \cos\phi\cos\theta
\end{bmatrix} \begin{bmatrix}
\dot{\phi} \\
\dot{\theta} \\
\dot{\psi}
\end{bmatrix} + \begin{bmatrix}
-\omega_o (\sin\psi/\cos\theta - \tan\theta\sin\theta\sin\psi) \\
-\omega_o (\cos\phi\cos\psi + \sin\phi\sin\theta\sin\psi) \\
\omega_o (\sin\phi\cos\psi - \cos\phi\sin\theta\sin\psi)
\end{bmatrix}.$$

(2.5)

where $\phi$, $\theta$ and $\psi$ denote the roll-, pitch- and yaw angles, respectively and $\omega_o$ is the angular velocity of the spacecraft orbit.

For small angles of $\phi$, $\theta$ and $\psi$, (2.5) can be simplified as:

$$\begin{bmatrix}
\omega_x \\
\omega_y \\
\omega_z
\end{bmatrix} = \begin{bmatrix}
\dot{\phi} - \omega_o \psi \\
\dot{\theta} - \omega_o \\
\dot{\psi} + \omega_o \phi
\end{bmatrix} = \begin{bmatrix}
\dot{\phi} \\
\dot{\theta} \\
\dot{\psi}
\end{bmatrix} + \begin{bmatrix}
0 & 0 & -\omega_o \\
0 & 0 & 0 \\
\omega_o & 0 & 0
\end{bmatrix} \begin{bmatrix}
\phi \\
\theta \\
\psi
\end{bmatrix} + \begin{bmatrix}
0 \\
\omega_o \\
0
\end{bmatrix}.$$

(2.6)

The matrix equivalent of (2.3) can be written as:
\[ \dot{\mathbf{r}}_{b,p} = \dot{\mathbf{r}}_{b,p,b} + \omega_{b,i} \times \mathbf{r}_{b,p,b}, \quad (2.7) \]

or:

\[ \dot{\mathbf{r}}_{b,p} = \mathbf{r}_{b,p,b} - \mathbf{r}_{b,p,b} \times \omega_{b,i}, \quad (2.8) \]

where \( \omega_{b,i} \) and \( \mathbf{r}_{b,p,b} \) are the skew symmetric matrices of \( \omega_{b,i} \) and \( \mathbf{r}_{b,p} \) in \( F_b \), respectively:

\[ \omega_{b,i} = \begin{bmatrix} 0 & -\omega_z & \omega_y \\ \omega_z & 0 & -\omega_x \\ -\omega_y & \omega_x & 0 \end{bmatrix}, \quad (2.9) \]

where \( \omega_x \), \( \omega_y \) and \( \omega_z \) denote the components of \( \omega_{b,i} \) in \( F_b \), see Eq. (2.4), and:

\[ \mathbf{r}_{b,p,b} = \begin{bmatrix} 0 & -r_z & r_y \\ r_z & 0 & -r_x \\ -r_y & r_x & 0 \end{bmatrix}, \quad (2.10) \]

where \( r_x \), \( r_y \) and \( r_z \) denote the components of \( \mathbf{r}_{b,p} \) in \( F_b \) (see Fig. 2.1). In the body fixed reference frame \( F_b \), \( \mathbf{r}_{b,p} \) is a time invariable vector since the main body is assumed to be rigid, i.e.:

\[ \dot{\mathbf{r}}_{b,p,b} = \mathbf{0}, \quad (2.11) \]

and therefore:

\[ \dot{\mathbf{r}}_{b,p} = \omega_{b,i} \times \mathbf{r}_{b,p,b}, \quad (2.12) \]
or:
\[ \dot{r}_{b,p} = -r_{b,p,b} \cdot \omega_{b,i} = \overset{T}{r}_{b,p,b} \cdot \omega_{b,i} \tag{2.13} \]

since \( r_{b,p,b} \) is a skew symmetric matrix.

Substituting (2.2) and (2.13) in (2.1), we obtain:
\[
dE_{kb} = \frac{1}{2} \left[ \overset{T}{r}_{i,b} \cdot \overset{.}{r}_{i,b} + \overset{T}{r}_{b,p} \cdot \overset{.}{r}_{b,p,b} - \overset{T}{r}_{b,p,b} \cdot \omega_{b,i} + \overset{T}{r}_{i,b} \cdot \overset{T}{r}_{b,p,b} \cdot \omega_{b,i} \right]. \tag{2.14} \]

The total kinetic energy of the rigid main body is obtained by integrating (2.14) over the entire volume of body as:
\[
E_{kb} = \int_{b} dE_{kb}
\]
\[
= \frac{1}{2} \left[ \overset{T}{r}_{i,b} \cdot \overset{.}{r}_{i,b} \int \text{dm} + \overset{T}{r}_{b,p,b} \cdot \overset{.}{r}_{b,p,b} \int \text{dm} \cdot \omega_{b,i} + \overset{T}{r}_{i,b} \cdot \overset{T}{r}_{b,p,b} \int \text{dm} \cdot \omega_{b,i} \right]. \tag{2.15} \]

The integral of the third term of (2.15) now is defined as the coupling matrix between the translational displacements and the rotational displacements of the rigid main body:
\[
Q_{b} = \int r_{b,p,b} \text{dm}, \tag{2.16} \]
where \( \mathbf{Q}_b \) is not necessarily to be zero since the origin of the rigid body fixed frame is not assumed as the mass center of the rigid main body. From Eq. (2.10) and (2.16) it follows that \( \mathbf{Q}_b \) has the following form:

\[
\mathbf{Q}_b = \begin{bmatrix}
0 & \int r_z \, dm & -\int r_y \, dm \\
-\int r_z \, dm & 0 & \int r_x \, dm \\
\int r_y \, dm & -\int r_x \, dm & 0 \\
\end{bmatrix}
\]

\[
= \begin{bmatrix}
0 & Q_z & -Q_y \\
-Q_z & 0 & Q_x \\
Q_y & -Q_x & 0 \\
\end{bmatrix}.
\]

The integral of the second term of Eq. (2.15) now is defined as the inertia matrix relative to the origin of the body fixed frame \( \mathbf{Q}_b \):

\[
\mathbf{I}_b = \int \mathbf{r}_{b,p,b}^T \mathbf{r}_{b,p,b} \, dm. \tag{2.17}
\]

From (2.10) and (2.17) it follows that \( \mathbf{I}_b \) has the following form:

\[
\mathbf{I}_b = \begin{bmatrix}
\int (r_y^2 + r_z^2) \, dm & -\int r_x r_y \, dm & -\int r_x r_z \, dm \\
-\int r_x r_y \, dm & \int (r_x^2 + r_z^2) \, dm & -\int r_y r_z \, dm \\
-\int r_x r_z \, dm & -\int r_y r_z \, dm & \int (r_x^2 + r_y^2) \, dm \\
\end{bmatrix}
\]

\[
= \begin{bmatrix}
I_x & -I_{xy} & -I_{xz} \\
-I_{xy} & I_y & -I_{yz} \\
-I_{xz} & -I_{yz} & I_z \\
\end{bmatrix}.
\]
Substituting (2.16) and (2.17) in (2.15), the kinetic energy of the rigid main body of the spacecraft can be obtained as:

\[ E_{kb} = \frac{1}{2} \dot{\Gamma}_{i,b} \cdot \dot{\Gamma}_{i,b} m_b + \frac{1}{2} \dot{\omega}_{b,i} \cdot I_b \cdot \dot{\omega}_{b,i} + \dot{\Gamma}_{i,b} \cdot \dot{Q}_b \cdot \omega_{b,i} \]  \hspace{1cm} (2.18)

where \( m_b \) is the total mass of the rigid main body.

The first term of Eq. (2.18) denotes the kinetic energy due to the translational displacements of the rigid main body, the second term of Eq. (2.18) denotes the kinetic energy due to the rotational displacements of the rigid main body and third term of Eq. (2.18) denotes coupling kinetic energy due to the translational and rotational displacements.

---

**Fig. 2.1** Representation of \( \Gamma_{b,p} \) in the body fixed reference frame \( F_b \).
2.2.2 The kinetic energy of the flexible substructures

In the following analysis, the kinetic energy of the flexible substructures will be obtained for the case where the rigid main body of the spacecraft is allowed to have translational and rotational motions. Moreover, the kinetic energy of the flexible substructures will be derived by using the finite element method.

As the first step in the application of the finite element method to the analysis of the flexible substructures, these substructures have to be idealized. In general, a flexible substructure, in the finite element method, is replaced by a finite collection of N structural elements. Each element is idealized as an elastic body. The elements are assumed to be connected to neighbours at discrete contact points, see Fig. 2.2. Each contact point is called a node.

Now, we define the flexible substructure local reference frame \( F_i \). As mentioned in Section 0.2, the local reference frame should be chosen in a suitable form for the analysis. In the finite element method, [8], [9] and [10], each element of the flexible substructure should have its own reference frame to analyse the deformation of the element in this frame. In this chapter we define that the local reference frame \( F_i \) is fixed in each undeformed element. For a particular flexible substructure shown in Fig. 2.3, assuming that there is a particle \( p \) in the flexible substructure with mass \( dm \), we can write the kinetic energy of the particle as:

\[
\delta E_{ka} = \frac{1}{2} \dot{R}_{i,p}^T \cdot \dot{R}_{i,p} \, dm, \tag{2.19}
\]

where \( \delta E_{ka} \) is the kinetic energy of the discussed particle and \( R_{i,p} \) is the vector from the origin \( O_i \) in the inertial reference frame \( F_i \) to the discussed particle \( p \) and `' ' means differentiation with respect to time.

From Fig. 2.3, the vector \( R_{i,p} \) can be decomposed as:

\[
R_{i,p} = R_{i,b} + R_{b,o} + R_{o,po} + R_{po,p}. \tag{2.20}
\]
where \( R_{b,0} \) is the vector from the origin \( O_b \) in the body fixed reference frame \( F_b \) to the origin \( O_1 \) in the flexible substructure local reference frame \( F_1 \). \( R_{0,po} \) is the vector from \( O_1 \) to any undeformed particle position \( p_0 \) and \( R_{po,p} \) is the vector from one undeformed particle \( p_0 \) to the corresponding deformed particle \( p \). This vector \( R_{po,p} \) will be called from now the distributed deformation vector.

Fig. 2.2 Some examples of flexible substructure idealizations and discretization as idealized elements.

The local reference frame \( F_1 \) may be moving relative to the body fixed reference frame \( F_b \), if the flexible substructure is allowed to have one or more kinematic degrees of freedom relative to the spacecraft rigid main body. For example, it may be desired that the spacecraft main body should point to the center of the earth while the flexible substructure should be
oriented towards the sun. Therefore, the transformation matrix $T_{b,1}$, see Section 0.2, should be a time variable matrix in general. A time variable matrix $T_{b,1}(t)$ will result in the total system to be time variable. It is assumed, however, that its variation is sufficiently slow to allow a quasi-stationary (i.e. time constant) approximation of the transformation matrix $T_{b,1}$ for relatively short time intervals. This approximation is reasonable since the data collection time interval for parameter identification is usually much shorter than the orbital period (see Chapter 1).

The body fixed reference frame, however, is still moving with respect to the inertial reference frame. Differentiation of (2.20) with respect to time in the inertial reference frame can be written as:

$$
\dot{R}_{1,b} = \dot{R}_{b,o} + R_{o,po} + R_{po,p}
$$

$$
= \dot{R}_{b,o,b} + R_{o,po,b} + R_{po,p,b} + \omega_{b,1} \times (R_{b,o,b} + R_{o,po,b} + R_{po,p,b}), \quad (2.21)
$$

where $\dot{R}_{b,o,b}$, $\dot{R}_{o,po,b}$ and $\dot{R}_{po,p,b}$ denote differentiations of $R_{b,o}$, $R_{o,po}$ and $R_{po,p}$ with respect to time in the body fixed reference frame $F_{b}$, and $R_{b,o,b}$, $R_{o,po,b}$ and $R_{po,p,b}$ denote $R_{b,o}$, $R_{o,po}$ and $R_{po,p}$ as represented in $F_{b}$.

Through the transformation matrix $T_{b,1}$, $R_{b,o,b}$, $R_{o,po,b}$ and $R_{po,p,b}$ can be represented in the flexible substructure local reference frame $F_{1}$:

$$
R_{b,o,1} + R_{o,po,1} + R_{po,p,1} = T_{b,1} \cdot (R_{b,o,b} + R_{o,po,b} + R_{po,p,b}), \quad (2.22)
$$

or:

$$
R_{b,o,b} + R_{o,po,b} + R_{po,p,b} = T_{b,1}^{-T} \cdot (R_{b,o,1} + R_{o,po,1} + R_{po,p,1}), \quad (2.23)
$$
where \( R_{b,0,1} \), \( R_{o,po,1} \) and \( R_{po,p,1} \) denote \( R_{b,0} \), \( R_{o,po} \) and \( R_{po,p} \) as represented in the flexible substructure local reference frame \( F_1 \).

It should be noted here, that \( T_{b,1}^{-1} \) is orthogonal so that \( T_{b,1}^{-1} = T_{b,1}^T \), see Eq. (2.29a).

![Diagram](image)

**Fig. 2.3** Example of a flexible substructure local reference frame \( F_1 \).

Referring to (2.23), (2.21) can also be written as:

\[
\dot{R}_{i,p} = \dot{T}_{i,b} + \dot{R}_{b,o} + \dot{R}_{o,po} + \dot{R}_{po,p}
\]

\[
= \dot{T}_{i,b} + T_{b,1}^T \cdot (\dot{R}_{b,o,1} + \dot{R}_{o,po,1} + \dot{R}_{po,p,1}) + \\
+ \omega_{b,i} \times [T_{b,1}^T \cdot (\dot{R}_{b,o,1} + \dot{R}_{o,po,1} + \dot{R}_{po,p,1})], \quad (2.24)
\]
where $\dot{R}_{b,0,1}$, $\dot{R}_{0,po,1}$ and $\dot{R}_{po,p,1}$ denote differentiations of $R_{b,0}$, $R_{0,po}$ and $R_{po,p}$ with respect to time in the flexible substructure local reference frame $F_1$.

Since $\dot{R}_{b,0}$ and $\dot{R}_{0,po}$ are constant vectors in the local reference frame, then

$$\dot{R}_{b,0,1} = \dot{R}_{0,po,1} = 0.$$  \hspace{1cm} (2.25)

When only small deformations of the flexible substructures are considered, the distributed deformation vector $R_{po,p,1}$ in (2.24) can be neglected as compared to $R_{b,0,1}$ and $R_{0,po,1}$ and (2.24) can be simplified as follows:

$$\ddot{R}_{i,p} = T_{i,b} + T_{b,1}^T \cdot R_{po,p,1} + \omega_{b,i} \times [T_{b,1}^T \cdot (R_{b,0,1} + R_{0,po,1})] .$$ \hspace{1cm} (2.26)

Using the laws of cross product of vectors, (2.26) can be written in its matrix equivalent form:

$$\ddot{R}_{i,p} = T_{i,b} + T_{b,1}^T \cdot R_{po,p,1} - T_{b,1}^T \cdot (R_{b,0,1} + R_{0,po,1}) \cdot T_{b,1} \cdot \omega_{b,i} ,$$ \hspace{1cm} (2.27)

where $R_{b,0,1}$ and $R_{0,po,1}$ are the skew symmetric matrix of $R_{b,0}$ and $R_{0,po}$ in the local reference frame $F_1$.

Substituting (2.27) in (2.19), the kinetic energy of the discussed particle $p$ in a flexible substructure can be written as:

$$dE_{ka} = \frac{1}{2} [T_{i,b} \cdot T_{i,b} + T_{po,p,1}^T \cdot R_{po,p,1} + \omega_{b,i} \cdot T_{b,1} \cdot (R_{b,0,1} + R_{0,po,1})^T \cdot (R_{b,0,1} + R_{0,po,1}) \cdot T_{b,1} \cdot \omega_{b,i} +$$
\[ + 2 \mathbf{\omega}_{b,i} \cdot \mathbf{T}_{b,i} \cdot (\mathbf{R}_{b,o,1} + \mathbf{R}_{o,po,1}) \cdot \mathbf{R}_{po,p,1} + \]
\[ + 2 \mathbf{\Gamma}_{b,1} \cdot \mathbf{T}_{b,1} \cdot (\mathbf{R}_{b,o,1} + \mathbf{R}_{o,po,1})^T \cdot \mathbf{T}_{b,1} \cdot \mathbf{\omega}_{b,i} + \]
\[ + 2 \mathbf{\Gamma}_{b,1} \cdot \mathbf{T}_{b,1} \cdot \mathbf{R}_{po,p,1} \] \text{dm} \quad (2.28) \]

In (2.28), the following identities have been used:
\[ \mathbf{T}_{b,1} \cdot \mathbf{T}_{b,1} = U_3 \quad (2.29a) \]

since \( \mathbf{T}_{b,1} \) is an orthogonal matrix, and:
\[ (\mathbf{R}_{b,po,1} + \mathbf{R}_{o,po,1})^T = - (\mathbf{R}_{b,o,1} + \mathbf{R}_{o,po,1}) \quad (2.29b) \]

since \( \mathbf{R}_{b,o,1} \) and \( \mathbf{R}_{o,po,1} \) are skew symmetric matrices.

The kinetic energy of the discussed element can be obtained by integrating (2.28) within one element, i.e.: 
\[ E_{k,a} = \int_{E_1} dE_{k,a} \]
\[ = \frac{1}{2} \mathbf{\Gamma}_{i,b} \cdot \mathbf{\Gamma}_{i,b} \cdot \int_{E_1} \text{dm} + \frac{1}{2} \int_{E_1} \mathbf{R}_{po,p,1}^T \cdot \mathbf{R}_{po,p,1} \text{dm} + \]
\[ + \frac{1}{2} \mathbf{\omega}_{b,i} \cdot \mathbf{T}_{b,1} \cdot \int_{E_1} (\mathbf{R}_{b,o,1} + \mathbf{R}_{o,po,1})^T (\mathbf{R}_{b,o,1} + \mathbf{R}_{o,po,1}) \text{dm} \mathbf{T}_{b,1} \cdot \mathbf{\omega}_{b,i} + \]
\[ + \mathbf{\Gamma}_{b,1} \cdot \mathbf{T}_{b,1} \cdot \int_{E_1} (\mathbf{R}_{b,o,1} + \mathbf{R}_{o,po,1}) \cdot \mathbf{R}_{po,p,1} \text{dm} + \]
\[ + \mathbf{e}_{i,b}^T \cdot \mathbf{z}_{b,l}^T \cdot \int_{E_i} \left( \mathbf{R}_{b,o,l} + \mathbf{R}_{o,po,l} \right)^T \, dm \, \mathbf{\omega}_{b,i} + \]

\[ + \mathbf{e}_{i,b}^T \cdot \mathbf{z}_{b,l}^T \cdot \int_{E_i} \mathbf{R}_{po,p,l} \, dm . \]

(2.30)

It should be noted that in (2.30) the integral \( \int \) means integration within \( E_i \) the whole element \( i \). The total kinetic energy of the flexible substructures is just the summation of all elementary kinetic energies. If the flexible substructures contain \( N \) elements, the total kinetic energy of the flexible substructures is:

\[ E_{ka} = \sum_{i=1}^{N} E_{ka_i} , \]

(2.31)

where \( E_{ka_i} \) is the kinetic energy of element \( i \).

Now, we turn to the finite element analysis for equation (2.30). In the finite element method, the distributed deformation displacement vector \( \mathbf{R}_{po,p,l} \) can be described by some combination of the shape function of an element and the nodal displacements of the element, and the static distribution \( \mathbf{R}_{o,po,l} \) can be represented by element nodal static locations. In the finite element method the integrals of the shape function of all element can be evaluated (see, e.g., [8], [9] and [10]), so that the quadratic functional (2.30) can be written in algebraic form.

To make the solution more clear, (2.30) will be discussed separately. First the integral of the fourth term of (2.30) can be decomposed as:

\[ \int_{E_i} \left( \mathbf{R}_{b,o,l} + \mathbf{R}_{o,po,l} \right) \cdot \mathbf{R}_{po,p,l} \, dm \]
since the integral is only within one element.

The displacement vectors as used in the finite element method are defined as follows.

\[ \mathbf{d}_{i,k} \] - displacement vector at node \( k \) in element \( i \). In general, it contains six components, i.e., three translational displacements and three rotational displacements.

\[ \mathbf{d}_i \] - assembled displacement vector of element \( i \). If element \( i \) contains \( n \) nodes, the displacement vector \( \mathbf{d}_i \) is:

\[
\mathbf{d}_i = [d_{i,1}^T \quad d_{i,2}^T \cdots d_{i,k-1}^T \quad d_{i,k}^T \quad d_{i,k+1}^T \cdots d_{i,n}^T]^T.
\]

\[ \mathbf{d}_k^* \] - displacement vector of the \( k^{th} \) node in the total displacement vector \( \mathbf{d} \) of the flexible substructures. In general, it consists of six components, i.e., three translational displacements and three rotational displacements.

\[ \mathbf{d} \] - total displacement vector of the flexible substructures. If the flexible substructures contain \( N_n \) nodes, then:

\[
\mathbf{d} = [(\mathbf{d}_1^*)^T \quad (\mathbf{d}_2^*)^T \cdots (\mathbf{d}_{k-1}^*)^T \quad (\mathbf{d}_k^*)^T \quad (\mathbf{d}_{k+1}^*)^T \cdots (\mathbf{d}_{N_n}^*)^T]^T.
\]

In the following discussion the distributed displacement \( R_{po,p,l} \) is to be a function vector of \( x, y \) and \( z \) which are the coordinates in the flexible substructure local reference frame \( F_1 \). Let:

\[
R_{po,p,l}(x, y, z) = F_1(x, y, z) \cdot c_i.
\]  

(2.33)
From the finite element method [8], [9] and [10], \( F_1(x, y, z) \), in general, is a \( 3 \times 6n \) matrix of assumed displacement functions since each node can be considered to have six degrees of freedom (three translational displacements and three rotational displacements). \( \mathbf{c}_1 \) is then a \( 6n \times 1 \) constant vector to be determined and \( n \) is the number of nodes in the element.

For the \( k \)th node of element \( i \), the displacement vector \( \mathbf{d}_{i,k} \) can be written as (see, e.g. [8] and [9]):

\[
\mathbf{d}_{i,k} = \frac{1}{2} \begin{bmatrix}
2U_3 \\
\mathbf{u}_3
\end{bmatrix}
\begin{bmatrix}
0 & -\partial/\partial z & \partial/\partial y \\
\partial/\partial z & 0 & -\partial/\partial x \\
-\partial/\partial y & \partial/\partial x & 0
\end{bmatrix}
\mathbf{c}_1 = \mathbf{Z}_{i,k} \cdot \mathbf{c}_1 , \quad (2.34)
\]

where \( U_3 \) is a \( 3 \times 3 \) unit matrix. All such \( \mathbf{d}_{i,k} \) are combined in \( \mathbf{d}_i \), i.e.:

\[
\mathbf{d}_i = \begin{bmatrix}
\mathbf{d}_{i,1} \\
\mathbf{d}_{i,2} \\
\vdots \\
\mathbf{d}_{i,n}
\end{bmatrix} = \begin{bmatrix}
\mathbf{Z}_{i,1} \\
\mathbf{Z}_{i,2} \\
\vdots \\
\mathbf{Z}_{i,n}
\end{bmatrix} \cdot \mathbf{c}_1 = \mathbf{Z}_i \cdot \mathbf{c}_1 . \quad (2.35)
\]

If \( \mathbf{Z}_i \) is of full rank, it follows from (2.35) that:

\[
\mathbf{c}_i = \mathbf{Z}_i^{-1} \cdot \mathbf{d}_i . \quad (2.36)
\]

Substituting (2.36) in (2.33), one can obtain:

\[
R_{po,p}(x, y, z) = F_1(x, y, z) \cdot \mathbf{Z}_i^{-1} \cdot \mathbf{d}_i
\]
\[ = \mathbf{C}_{z_1}(x, y, z) \cdot \mathbf{d}_i, \]  
(2.37)

where:

\[ \mathbf{C}_{z_1}(x, y, z) = \mathbf{F}_{z_1}(x, y, z) \cdot \mathbf{Z}_{z_1}^{-1}. \]  
(2.38)

In the finite element method, \( \mathbf{C}_{z_1}(x, y, z) \) is called the shape function matrix of element \( i \).

The nodal displacement vector \( \mathbf{d}_i \) can also be arranged as follows:

\[
\mathbf{d}_i = \begin{bmatrix} d_{i,t} \\ \vdots \\ d_{i,r} \end{bmatrix}.
\]  
(2.39)

where \( d_{i,t} \) contains all the nodal translational displacements of the \( i \)th element and \( d_{i,r} \) contains all the nodal rotational displacements of the \( i \)th element. With this arrangement the matrix \( \mathbf{C}_{z_1} \) can be partitioned as:

\[
\mathbf{C}_{z_1} = \begin{pmatrix} \mathbf{C}_{z_1,t} & \mathbf{C}_{z_1,r} \end{pmatrix},
\]  
(2.40)

where \( \mathbf{C}_{z_1,t} \) is the shape function matrix with respect to the nodal translational displacements \( d_{i,t} \) and \( \mathbf{C}_{z_1,r} \) is the shape function matrix with respect to the nodal rotational displacements \( d_{i,r} \). Substituting (2.39) and (2.40) in (2.37), it follows that:

\[
\mathbf{R}_{p_0,p_1}(x, y, z) = \begin{pmatrix} \mathbf{C}_{z_1,t} & \mathbf{C}_{z_1,r} \end{pmatrix} \begin{bmatrix} d_{i,t} \\ \vdots \\ d_{i,r} \end{bmatrix}.
\]  
(2.41)
Now we define the following matrices:

\[
M_{z_i} = \int_{E_i} \mathbf{C}_i^T \cdot \mathbf{C}_i \, dm
\]

\[
= Z_{z_i}^{-T} \cdot \int_{E_i} \mathbf{F}_i^T \cdot \mathbf{F}_i \, dm \cdot Z_{z_i}^{-1}, \tag{2.42}
\]

\[
M_{z_i,t} = \int_{E_i} \mathbf{C}_i^T \cdot \mathbf{C}_i \, dm, \tag{2.43}
\]

\[
M_{z_i,r} = \int_{E_i} \mathbf{C}_i^T \cdot \mathbf{C}_i \, dm. \tag{2.44}
\]

where \( M_{z_i} \) is called the total mass matrix with respect to all nodal displacements of the \( i \)th element, \( M_{z_i,t} \) is the mass matrix related to the nodal translational displacements of the \( i \)th element and \( M_{z_i,r} \) is the mass (inertia) matrix related to the nodal rotational displacements of the \( i \)th element. Substituting (2.37) in the right hand side of (2.32), one can write for the \( i \)th element:

\[
R_{b,0,1} \cdot \int_{E_i} R_{p0,p,1} \, dm + \int_{E_i} R_{0,p0,1} \cdot R_{p0,p,1} \, dm
\]

\[
= R_{b,0,1} \cdot \int_{E_i} \mathbf{C}_i \, dm \cdot \dot{\mathbf{d}}_i + \int_{E_i} R_{0,p0,1} \cdot \mathbf{C}_i \, dm \cdot \dot{\mathbf{d}}_i
\]

\[
= (R_{b,0,1} \cdot \mathbf{W}_z + \mathbf{V}_z) \cdot \dot{\mathbf{d}}_i
\]
\[ = \mathbf{A}_1 \cdot \mathbf{d}_1 \quad , \]  
(2.45)

where:

\[ \mathbf{W}_1 = \int \mathbf{C}_{\mathbf{C}_1} \, \mathbf{d}m \quad , \]  
(2.46)

\[ \mathbf{V}_1 = \int \mathbf{R}_{\mathbf{R}_0, \mathbf{R}_p, \mathbf{R}_L} \cdot \mathbf{C}_{\mathbf{C}_1} \, \mathbf{d}m \quad , \]  
(2.47)

\[ \mathbf{A}_1 = \mathbf{R}_{\mathbf{R}_0, \mathbf{R}_p, \mathbf{R}_L} \cdot \mathbf{W}_1 + \mathbf{V}_1 \quad . \]  
(2.48)

Using (2.37) and (2.42), the second term of (2.30) can be written as:

\[
\int \mathbf{R}_{\mathbf{R}_p, \mathbf{R}_L}^T \cdot \mathbf{R}_{\mathbf{R}_p, \mathbf{R}_L} \, \mathbf{d}m = \mathbf{d}_1^T \cdot \int \mathbf{C}_{\mathbf{C}_1}^T \cdot \mathbf{C}_{\mathbf{C}_1} \, \mathbf{d}m \cdot \mathbf{d}_1
\]

\[ = \mathbf{d}_1^T \cdot \mathbf{M}_{\mathbf{M}_1} \cdot \mathbf{d}_1 \quad . \]  
(2.49)

The integral of the third term of (2.30) is defined as:

\[ \mathbf{I}_1 = \int \left( \mathbf{R}_{\mathbf{R}_0, \mathbf{R}_p, \mathbf{R}_L} + \mathbf{R}_{\mathbf{R}_0, \mathbf{R}_p, \mathbf{R}_L} \right)^T \cdot \left( \mathbf{R}_{\mathbf{R}_0, \mathbf{R}_p, \mathbf{R}_L} + \mathbf{R}_{\mathbf{R}_0, \mathbf{R}_p, \mathbf{R}_L} \right) \, \mathbf{d}m
\]

\[ = \mathbf{R}_{\mathbf{R}_0, \mathbf{R}_p, \mathbf{R}_L}^T \cdot \mathbf{R}_{\mathbf{R}_0, \mathbf{R}_p, \mathbf{R}_L} \cdot \mathbf{m}_1 + \mathbf{R}_{\mathbf{R}_0, \mathbf{R}_p, \mathbf{R}_L}^T \int \mathbf{R}_{\mathbf{R}_0, \mathbf{R}_p, \mathbf{R}_L} \, \mathbf{d}m +
\]

\[ + \left( \mathbf{R}_{\mathbf{R}_0, \mathbf{R}_p, \mathbf{R}_L}^T \cdot \int \mathbf{R}_{\mathbf{R}_0, \mathbf{R}_p, \mathbf{R}_L} \, \mathbf{d}m \right)^T + \int \mathbf{R}_{\mathbf{R}_0, \mathbf{R}_p, \mathbf{R}_L}^T \cdot \mathbf{R}_{\mathbf{R}_0, \mathbf{R}_p, \mathbf{R}_L} \, \mathbf{d}m \quad , \]  
(2.50)

where \( \mathbf{m}_1 \) is the total mass of the discussed element.
From the finite element method (see, e.g., [8], [9] and [10]), it is usually assumed that the mass density within the chosen element is uniform, therefore, (2.50) can also be written as:

\[ I_{i} = \rho_{i} v_{i} R_{b,o,l}^{T} + L_{b,o,l}^{T} + I_{E_{i}}^{T}, \quad (2.51) \]

where \( \rho_{i} \) is the mass density of the \( i \)th element, \( v_{i} \) is the volume of the \( i \)th element and:

\[ I_{E_{i}} = \rho_{i} \left[ R_{b,o,l}^{T} + \int_{E_{i}} R_{o,po,l}^{T} \, dv + \left( R_{o,po,l}^{T} + \int_{E_{i}} R_{o,po,l}^{T} \, dv \right)^{T} \right] + \int_{E_{i}} R_{o,po,l}^{T} \, dv \]. \quad (2.52) \]

From (2.45), the last term of (2.30) can be written as:

\[ \dot{r}_{i,b} \cdot T_{b,l}^{T} + \int_{E_{i}} \dot{r}_{o,po,p} \cdot T_{b,l}^{T} \cdot W_{i} \cdot \dot{d} \quad (2.53) \]

Finally, we define the integral of the fifth term of (2.30) as:

\[ \int_{E_{i}} \left( R_{b,o,l} + R_{o,po,l} \right)^{T} \, dm = Q_{i} \quad (2.54) \]

Substituting (2.45), (2.49), (2.50), (2.53) and (2.54) into (2.30), the kinetic energy of the discussed element can be written as:

\[ E_{k,i} = \frac{1}{2} \dot{r}_{i,b} \cdot \dot{r}_{i,b} \cdot m_{i} + \frac{1}{2} \dot{d} \cdot M_{i} \cdot \dot{d} + \]
\[ + \frac{1}{2} \omega_{b,i}^T \cdot \frac{b}{b,i} \cdot I_{i} \cdot \frac{b}{b,1} \cdot \omega_{b,i}^T + \omega_{b,i}^T \cdot T_{b,1}^T \cdot A_i \cdot d_i + \]

\[ + \frac{T_{i,b}^T \cdot \frac{b}{b,1} \cdot \omega_{b,i} +}{+ \frac{T_{i,b}^T \cdot \frac{b}{b,1} \cdot \omega_{b,i} +}{+ \frac{T_{b,1}^T \cdot \omega_{b,i} +}{+ \frac{T_{b,1}^T \cdot \omega_{b,i} +}{+ \frac{T_{b,1}^T \cdot \omega_{b,i} +}{+ \frac{T_{b,1}^T \cdot \omega_{b,i} +}{(2.55)}} \]

Since the relation between the \textit{i}^{th} element displacement vector \(d_i\) and the total structure displacement vector \(\delta\) can be written as:

\[ d_i = p_{i,1}^T \delta \quad , \quad (2.56) \]

where \(p_{i,1}^T\) is called the assembling matrix.

The assembling matrix \(p_{i,1}^T\) depends on the element displacement vector \(d_i\) and the total structure displacement vector \(\delta\) as shown in Eq. (2.56). Since \(d_i\) consists of the nodal displacement vector as well, it is easy to find the relation between \(d_i\) and \(\delta\). For example, assuming that there are four nodes in the \textit{i}^{th} element, i.e.:

\[ d_i = [d_{i,1}^T \ d_{i,2}^T \ d_{i,3}^T \ d_{i,4}^T]^T \]

and each node has six degrees of freedom, i.e.:

\[ d_{i,k} = [d_{i,k,1} \ d_{i,k,2} \ ... \ d_{i,k,6}]^T \]

and assuming that the corresponding nodal displacement vectors are \(d_{a}^*, d_{b}^*, d_{c}^*\) and \(d_{d}^*\) in the total displacement vector \(\delta\), i.e.:
\[ d = [(d_1^*)^T (d_2^*)^T \ldots (d_a^*)^T (d_b^*)^T \ldots (d_{k-1}^*)^T (d_k^*)^T (d_{k+1}^*)^T \ldots \\
\ldots (d_d^*)^T (d_e^*)^T \ldots (d_{Nn}^*)^T]^T, \]

the assembling matrix \( P_{1i} \) of the element \( i \) can then be written as:

\[
P_{1i} = \begin{bmatrix}
0 & 0 & \ldots & U_{6} & 0 & \ldots & 0 & 0 & \ldots & 0 \\
0 & 0 & \ldots & 0 & U_{6} & \ldots & 0 & 0 & \ldots & 0 \\
0 & 0 & \ldots & 0 & 0 & \ldots & 0 & U_{6} & 0 & \ldots & 0 \\
0 & 0 & \ldots & 0 & 0 & \ldots & 0 & 0 & \ldots & U_{6} & 0 \\
0 & 0 & \ldots & 0 & 0 & \ldots & 0 & 0 & \ldots & 0 & U_{6} \\
0 & 0 & \ldots & 0 & 0 & \ldots & 0 & 0 & \ldots & 0 & 0 \\
0 & 0 & \ldots & 0 & 0 & \ldots & 0 & 0 & \ldots & 0 & 0 \\
\end{bmatrix}.
\]

1 2 \ldots a b \ldots k-1 k k+1 \ldots c d \ldots Nn
(sequence number of node)

Moreover, the assembling matrix \( P_{1i} \) has the following property:

\[ P_{1i} \cdot P_{1i}^T = U_{6n}. \tag{2.57} \]

where \( n \) is the number of nodes in the element \( i \).

Using (2.55) and (2.56), the total kinetic energy of the flexible substructures can be derived from (2.31) as:

\[
E_{ka} = \frac{1}{2} \sum_{i=1}^{N} \left[ \sum_{m=1}^{m_i} \right]^T \left[ \sum_{i=1}^{N} P_{1i}^T \cdot M_{ii} \cdot P_{1i} \right] \cdot d + \\
+ \frac{1}{2} \omega_{b,i} \cdot \left[ \sum_{i=1}^{N} I_{ii}^T \cdot I_{ii} \right] \cdot \omega_{b,i} + \omega_{b,i}^T \cdot \left[ \sum_{i=1}^{N} I_{ii}^T \cdot A_{ii} \cdot P_{1i} \right] \cdot d + \\
+ \sum_{i=1}^{N} \left[ \sum_{i=1}^{N} I_{ii}^T \cdot Q_{ii} \cdot I_{ii} \right] \cdot \omega_{b,i} + \sum_{i=1}^{N} \left[ \sum_{i=1}^{N} I_{ii}^T \cdot W_{ii} \cdot P_{1i} \right] \cdot d. \tag{2.58}
\]
Now, we define:

\[ m_a = \sum_{i=1}^{N} m_i \]  

as the total mass of the flexible substructures,

\[ M = \sum_{i=1}^{N} P_{q1}^T \cdot M_i \cdot P_{q1} \]  

as the total mass matrix of the flexible substructures, see Eq. (2.42),

\[ I_{qa} = \sum_{i=1}^{N} T_{q1}^T \cdot I_i \cdot T_{q1} \]  

as the total inertia matrix of the flexible substructures, see Eq. (2.51),

\[ A = \sum_{i=1}^{N} T_{q1}^T \cdot A_i \cdot P_{q1} \]  

as the total coupling matrix for the rotational displacements of the rigid main body and the flexible substructure displacements, see Eq. (2.48),

\[ W = \sum_{i=1}^{N} T_{q1}^T \cdot W_i \cdot P_{q1} \]  

as the total coupling matrix for the translational displacements of the rigid main body and the flexible substructure displacements, see Eq. (2.46), and

\[ Q_{qa} = \sum_{i=1}^{N} T_{q1}^T \cdot Q_i \cdot T_{q1} \]  

(2.64)
as the total coupling matrix for the translational displacements of the rigid main body and the rotational displacements of the rigid main body, see Eq. (2.54).

Substituting all the defined matrices above in (2.58), we obtain:

\[
E_{ka} = \frac{1}{2} \dot{\mathbf{r}}_{i,b} \cdot \mathbf{r}_{i,b} \mathbf{m}_a + \frac{1}{2} \dot{\mathbf{d}} \cdot \mathbf{M} \cdot \mathbf{d} + \frac{1}{2} \omega_{b,i} \cdot \mathbf{I}_a \cdot \omega_{b,i} + \\
+ \omega_{b,i} \cdot \mathbf{A} \cdot \mathbf{d} + \mathbf{T}_{i,b} \cdot \mathbf{W} \cdot \mathbf{d} + \frac{1}{2} \mathbf{Q}_b \cdot \omega_{b,i}.
\]  

(2.65)

As an explanation of Eq. (2.65), it can be expressed, that the first term denotes the kinetic energy due to the translational displacements of the undeformed flexible substructures, the second term denotes the kinetic energy due to the rotational displacements of the undeformed flexible substructures and the third term denotes the kinetic energy due to the deformations of the flexible substructures, and it can also be found, that the last three terms denote the coupling kinetic energies due to the translational displacements of the undeformed flexible substructures, the rotational displacements of the undeformed flexible substructures and the deformations of the flexible substructures.

It is noted here, that in (2.55) to (2.64) we have used \( T_{i1} \) instead of \( T_{b1} \) since we can, in general, assume that the local reference frame may be different for every element. In a particular case, for example, a spacecraft with a rigid central body and two symmetrical flexible plate solar panels, the number of the transformation matrices \( T_{i1} \) can be reduced to two for the two solar panels, see Section 2.5.

2.2.3 The total kinetic energy of the spacecraft

The total kinetic energy of the flexible spacecraft can be obtained by taking the sum of the kinetic energies of the rigid main body and the kinetic energy of the flexible substructures, i.e.:

\[
E_k = E_{kb} + E_{ka}.
\]  

(2.66)
Refering to (2.18) and (2.65), (2.66) can be written as:

\[
E_K = \frac{1}{2} \Gamma_{i,b}^T \cdot \Gamma_{i,b} \cdot m + \frac{1}{2} \dot{\Gamma}_{i,b}^T \cdot M \cdot \dot{\Gamma}_{i,b} + \frac{1}{2} \omega_{b,i}^T \cdot I_{\omega} \cdot \omega_{b,i} + \\
+ \omega_{b,i}^T \cdot A \cdot d + \Gamma_{i,b}^T \cdot \Omega \cdot d + \Gamma_{i,b}^T \cdot Q \cdot \omega_{b,i},
\]  

(2.67)

where:

\[
I = I_b + I_a
\]

as the total inertia matrix of the spacecraft,

\[
m = m_b + m_a
\]

as the total mass of the spacecraft, and:

\[
Q = Q_b + Q_a
\]

as the total coupling matrix between the translational and the rotational displacements of the undeformed spacecraft.

2.3 The Potential Energy of the Flexible Spacecraft

In principle, the potential energy of a flexible spacecraft consists of the potential energy of the undeformed spacecraft and the potential energy due to the deformation of the flexible substructures. In this section these two parts of the potential energy will be discussed separately.

2.3.1 The potential energy of the undeformed spacecraft

The potential energy of an undeformed spacecraft can be considered as the potential energy of a rigid spacecraft. In the case of space motion, the motion of a spacecraft relative to the earth or other planets should be treated
as the two-body problem. In this study, the spacecraft motion is related to
the earth, so that we only discuss the potential energy of the spacecraft
relative to the earth.
When the spacecraft motion is related to the earth, the spacecraft motion
can be considered as a particle motion relative to the earth since the mass
of the earth is much larger than the mass of the spacecraft. The motion of
the spacecraft then takes place about the the earth.
If a force vector \( \mathbf{F}_r \) acting on a particle moves through a distance vector
\( \mathbf{dr} \), the work done is equal to:

\[
dW = \mathbf{F}_r^T \cdot \mathbf{dr} .
\]  

(2.68a)

On the other hand, when the force \( \mathbf{F}_r \) is conservative, the above equation can
be written as:

\[
dW = \mathbf{F}_r^T \cdot \mathbf{dr} = - dE_{pr} ,
\]  

(2.68b)

which expresses the conservative force in terms of the rigid spacecraft
potential energy \( E_{pr} \).
Eq. (2.68b) can also be written as:

\[
W = - E_{pr}(r) ,
\]  

(2.68c)

The potential energy \( E_{pr}(r) \) now is a function of the rigid spacecraft dis-
placement vector \( r \).
The vector \( r \) is the vector from the center of the earth to the mass center
of the undeformed spacecraft and it can be decomposed as:

\[
r = r_{e,i} + r_{i,b} ,
\]  

(2.68d)
where \( r_{e,i} \) denotes the vector from the mass center of the earth to the origin of the orbital reference frame. When discussed the spacecraft nominal orbit is circular, vector \( r_{e,i} \) should be with a constant length. Therefore, the potential energy of the undeformed spacecraft can be written as a function of the translational displacement vector \( r_{i,b} \) of the undeformed spacecraft:

\[
E_{\text{pr}} = E_{\text{pr}}(r_{i,b}) \quad (2.68e)
\]

In this study the orbital motion of the spacecraft will not be discussed, so that the influences of the potential energy of the undeformed spacecraft will be represented by external forces acting on the spacecraft. Detailed discussions of this conversion is given in Appendix A.

2.3.2 The potential energy due to the deformations of the flexible substructures

If the flexible spacecraft consists of a rigid main body and some flexible substructures, the potential energy due to the deformations of the flexible substructures is equal to the sum of the strain energy of the flexible substructures and the potential energy due to the external forces acting on the substructures with minus sign.

The potential energy of a particular element in a flexible substructure can be found from [8], [9] and [10] as follows:

\[
E_{pa_i} = \frac{1}{2} \int_{E_i} ^{T} \sigma^T_i \cdot \varepsilon_i \, dv - \int_{E_i} ^{T} R_{p_0, p, l}^T \cdot F_r \, dv \quad (2.69)
\]

where \( \sigma_i \) is the vector of stress, \( \varepsilon_i \) is the vector of strain and \( F_r \) is so-called the distributed external force vector acting on the element. The strains \( \varepsilon_i(x, y, z) \) at any point \((x, y, z)\) within the element are now related to the distributed displacements \( R_{p_0, p, l}(x, y, z) \) at that point and hence to the nodal displacement vector \( d_i \).
The first term of Eq. (2.69) is the strain energy of the deformed element and the second term of Eq. (2.69) is the potential energy due to external forces acting on the element.

The strain at any point within an element may be obtained from the chosen displacement function matrix $\mathbf{F}_1(x, y, z)$ described in Section 2.2 by differentiation, the exact form of the differentiation depending upon the type of problem being considered. For example, for a plane elasticity problem the strains correspond to the first derivative of the displacements, while for a flexural problem the strains are associated with the curvature of the element and correspond to the second derivative of the displacements (see, e.g., [8], [9] and [10]). In general:

$$\varepsilon_i = \mathbf{B}_i \cdot \mathbf{R}_{p_0,p,l},$$  \hspace{1cm} (2.70)

where $\mathbf{B}_i$ is an operator matrix which contains first or second order derivative operators.

The exact form of (2.70) for any particular class of problem may be obtained from the theory of elasticity. Using the expression for $\mathbf{R}_{p_0,p,l}$ given by (2.37) and noting that both $Z_i^{-1}$ and $d_i$ are independent of $x$, $y$ and $z$, the strain vector $\varepsilon$ is given by:

$$\varepsilon_i(x, y, z) = \mathbf{B}_i \cdot \mathbf{F}_i \cdot Z_i^{-1} \cdot d_i$$

$$= G_i \cdot Z_i^{-1} \cdot d_i,$$  \hspace{1cm} (2.70a)

where:

$$G_i = \mathbf{B}_i \cdot \mathbf{F}_i.$$

In general, matrix $G_i$ contains terms in $x$, $y$ and $z$. 
Since a relation between the internal strains and nodal displacements $d_i$ is already known in (2.70a), the internal stresses $g_i(x, y, z)$ can be related to the nodal displacements from the relation between $e_i$ and $q_i$. Clearly, in this step the elastic properties of the element have to be taken into consideration.

In general, the relation between internal strains and stresses is given (see, [8] and [9]) by:

$$ g_i(x, y, z) = R_i \cdot e_i(x, y, z) , \quad (2.71) $$

where $R_i$ is termed the elasticity matrix. It contains the elastic properties of the $i$-th element, i.e., quantities such as the Young's modulus of elasticity $E$ and Poisson's ratio $v$. Referring to (2.70a), (2.71) can be written as:

$$ g_i(x, y, z) = R_i \cdot G_i \cdot Z_i^{-1} \cdot d_i \quad (2.72) $$

Substituting (2.71) and (2.72) in (2.69), the potential energy of the element can be obtained as:

$$ E_{p_i} = \frac{1}{2} d_i^T \cdot Z_i^{-T} \cdot \int_{E_i} G_i^T \cdot R_i \cdot G_i \cdot d\mathbf{v} \cdot Z_i^{-1} \cdot d_i - d_i^T \cdot \int_{E_i} G_i^T \cdot F_p \cdot d\mathbf{v} \quad (2.73) $$

Now we define:

$$ K_i = Z_i^{-T} \cdot \int_{E_i} G_i^T \cdot R_i \cdot G_i \cdot d\mathbf{v} \cdot Z_i^{-1} \quad , \quad (2.74a) $$

as the stiffness matrix of element $i$, and:
\[ F_i = \int_{E_1} G_i^T \cdot F_i \, dv \]  
(2.74b)

as the so-called discrete external forces acting on the nodes of element \( i \). Therefore, the potential energy of the deformed element can be written as:

\[ E_{pa_i} = \frac{1}{2} d_{i1}^T \cdot K_{i1} \cdot d_{i1} - d_{i1}^T \cdot F_i \]  
(2.75)

Considering the relationship between \( d_{i1} \) and the total nodal displacement vector \( d \) in (2.56), the total potential energy of the flexible substructures due to structural deformations is:

\[ E_{pa} = \frac{1}{2} d^T \cdot \left( \sum_{i=1}^{N} P_i^T \cdot K_{i1} \cdot P_i \right) \cdot d - d^T \cdot \left( \sum_{i=1}^{N} P_i^T \cdot F_i \right) \]  
(2.76)

Let the total stiffness matrix be:

\[ K = \sum_{i=1}^{N} P_i^T \cdot K_{i1} \cdot P_i \]  
(2.77a)

and the total external forces acting on the flexible substructures be:

\[ F_a = \sum_{i=1}^{N} P_i^T \cdot F_i \]  
(2.77b)

(2.76) can then be written as:

\[ E_{pa} = \frac{1}{2} d^T \cdot K \cdot d - d^T \cdot F_a \]  
(2.78)
2.3.3 The total potential energy of the spacecraft

The total potential energy of a flexible spacecraft can be obtained from the potential energy of the undeformed spacecraft and the potential energy due to the structural deformation as mentioned in the beginning of this section as:

\[ E_p = E_{p_{\text{r}}} + E_{p_{\text{a}}} \]  \hspace{1cm} (2.79a)

Referring to Eqs. (2.68e) and (2.78), Eq. (2.79a) can be written as:

\[ E_p = E_{p_{\text{r}}}(r_{1,b}) + \frac{1}{2} d^T K d - d^T F_a \]  \hspace{1cm} (2.79b)

2.4 A Set of General Linear Differential Equations of Motion of Flexible Spacecraft

To derive the equations of motion of a flexible spacecraft from its kinetic energy and potential energy, Lagrange's formulations of the equations of motion can be used. The Lagrangian \( L \) has the following form:

\[ L = E_k - E_p \]  \hspace{1cm} (2.80)

Substituting (2.67) and (2.79b) in (2.80), the Lagrangian can be expressed by the displacements and velocities of the spacecraft as follows:

\[ L = \frac{1}{2} \dot{r}^T m \dot{r} + \frac{1}{2} \dot{d}^T M \dot{d} + \frac{1}{2} \dot{\omega}^T I \dot{\omega} + \dot{\omega}^T A d + \dot{r}^T W d + \dot{r}^T Q \omega + \]

\[ - E_{p_{\text{r}}}(r) - \frac{1}{2} d^T K d + d^T F_a \]  \hspace{1cm} (2.81)

In Eq. (2.81) \( r = r_{1,b} \) and \( \omega = \omega_{b,1} \) have been used for simplicity.

Since the displacements occurring in (2.81) are represented in different reference frames, the Lagrange's equations have certain nonlinear terms.
Detailed conversion of this type of Lagrange's equations is given in Appendix A. It can be recognized as so-called the Lagrange's equations for quasi-coordinates, see [29] and [30]. As mentioned in Chapter 1, only small position and attitude errors of the rigid main body of the spacecraft and small deformations of the flexible substructures are considered in this study, so that the equations of motion may be linearized and the extra nonlinear terms will disappear, see also Appendix A. Therefore, a set of linear equations of motion of the flexible spacecraft can be derived as follows:

\[
m \ddot{r} + Q (\ddot{\theta} + \omega_o \dot{\theta}) + W \ddot{d} = F_b , \tag{2.82}
\]

\[
Q^T \dddot{r} + I (\dddot{\theta} + \omega_o \dot{\theta}) + A \dddot{d} = T_b , \tag{2.83}
\]

\[
W^T \dddot{r} + A^T (\dddot{\theta} + \omega_o \dot{\theta}) + M \dddot{d} + D \dot{d} + K d = F_a , \tag{2.84}
\]

where:

\[
\ddot{\theta} = [\ddot{\phi} \dddot{\theta} \dddot{\psi}]^T , \tag{2.85}
\]

\[
\dot{\theta} = [\dot{\phi} \dot{\theta} \dot{\psi}]^T , \tag{2.86}
\]

and:

\[
\omega_o = \begin{bmatrix}
0 & 0 & -\omega_o \\
0 & 0 & 0 \\
\omega_o & 0 & 0
\end{bmatrix} . \tag{2.87}
\]

In Eqs. (2.82), (2.83) and (2.84) $F_b$, $F_a$ and $T_b$ are the force and torque vectors acting on the rigid main body and every node of the flexible substructures of the spacecraft, see Appendix A.
So-called proportional damping, see [31], leads to the simplest form of the damping matrix $\mathbf{D}$. The damping matrix $\mathbf{D}$ in proportional damping is the linear combination of the system mass matrix $\mathbf{M}$ and the stiffness matrix $\mathbf{K}$, i.e.:

$$
\mathbf{D} = \alpha \mathbf{M} + \beta \mathbf{K},
$$

(2.88)

where for a general flexible structure $\alpha$ and $\beta$ are the constants to be determined from two given damping ratios that correspond to two unequal natural frequencies of vibration [31]. If two natural frequencies $\omega_i$ ($i = 1, 2$) and two damping ratios $\zeta_i$ ($i = 1, 2$) are known, $\alpha$ and $\beta$ can be obtained from:

$$
\alpha + \beta \frac{\omega^2_i}{\omega_i} = 2 \omega_i \zeta_i \quad (i = 1, 2),
$$

(2.89a)

and all damping ratios can be obtained from:

$$
\zeta_i = \frac{\alpha + \beta \frac{\omega^2_i}{\omega_i}}{2 \omega_i} \quad (i = 1, 2, \ldots, m),
$$

(2.89b)

where $m$ is the total modes of the flexible structure. However, as these damping ratios and natural frequencies may not be assumed to be known prior to identification, (2.89a) and (2.89b) can not be used. Consequently, $\alpha$ and $\beta$ must be considered as the unknown parameters to be identified. The reason to choose the proportional damping in our study is that only two parameters in the damping matrix need to be estimated rather than all the elements of matrix $\mathbf{D}$, see Section 2.6.

In matrix form, (2.82), (2.83) and (2.84) can be written as:
\[
\begin{bmatrix}
\begin{bmatrix}
mU_3 & Q & W \\
Q^T & I & A \\
W^T & A^T & M
\end{bmatrix} & \begin{bmatrix}
0 & Q \omega_0 & 0_{3x3} \\
0 & I \omega_0 & 0_{3x3} \\
0 & 0_{3x3} & D
\end{bmatrix} & \begin{bmatrix}
0 & 0 & 0_{3x3} \\
0 & 0 & 0_{3x3} \\
0 & 0 & K
\end{bmatrix}
\end{bmatrix} \begin{bmatrix}
\theta \\
\theta \\
\theta
\end{bmatrix} = \begin{bmatrix}
F_b \\
F_b \\
F_a
\end{bmatrix}
\]

(2.90)

Equation (2.90) can also be written in the first order differential equation form as follows:

\[
\begin{bmatrix}
F \\
F \\
F
\end{bmatrix} \begin{bmatrix}
\ddot{x}_0 \\
\dot{x}_0 \\
x_0
\end{bmatrix} = \begin{bmatrix}
A_0 & \begin{bmatrix}
B_0 & 0_{6+m}
\end{bmatrix}
\end{bmatrix} \begin{bmatrix}
\dot{x}_0 \\
x_0
\end{bmatrix} + \begin{bmatrix}
0_{6+m} & 0_{6+m}
\end{bmatrix} u_0 ,
\]

(2.91)

where:

\[
x_0 = \begin{bmatrix}
\begin{bmatrix}
\dot{r} & \dot{\theta} & \dot{d} & r^T & \theta^T & d^T
\end{bmatrix}^T
\end{bmatrix}
\]

\[
u_0 = \begin{bmatrix}
F_b^T & T_b^T & F_a^T
\end{bmatrix}^T
\]

\[
F_0 = \begin{bmatrix}
mU_3 & Q & W & 0_{3x(6+m)} \\
Q^T & I & A & 0_{3x(6+m)} \\
W^T & A^T & M & 0_{m(6+m)} \\
0_{(6+m)x3} & 0_{(6+m)x3} & 0_{(6+m)xm} & U_{6+m}
\end{bmatrix}
\]

\[
B_0 = \begin{bmatrix}
0_{6+m} & 0_{6+m}
\end{bmatrix}^T
\]
2.5 A Particular Dynamical Model of a Spacecraft
with Two Symmetrical Solar Panels

In the previous sections of this chapter, the general representation of flexible spacecraft dynamics has been obtained. For simulation, the matrices of the model (e.g., mass matrix \( M \), stiffness matrix \( K \) and the damping matrix \( D \)) should be discussed in detail. In this section, a specific flexible spacecraft with a rigid central body and two symmetrical plate solar panels is chosen, for which these matrices will be evaluated.

2.5.1 Model idealization

In this study the data to be used in experiment are chosen, to a certain extent arbitrarily as follows.

1. Geostationary orbit, the z axis of the body fixed frame should point to the center of the earth, the y axis is normal to the orbital plane and the x axis should point to the orbit direction when there are no attitude errors.

2. Two large symmetrical plate solar panels are attached and oriented towards the sun. In this case the inertia matrix \( I_a \) of the flexible solar panels should vary with time for a whole orbital period. For relatively short time intervals, this matrix will be considered as a constant matrix.
3. The translational displacements of the rigid central body are assumed to be zero, so that the origin of the inertial reference frame coincides with the origin of the body fixed frame.

The discussed flexible spacecraft is shown in Fig. 2.4.

Using these definitions, the transformation matrices between the body fixed frame and the flexible solar panel local frames are reduced to two for the two solar panels. For the right hand side solar panel, the transformation matrix is (see Fig. 2.4):

\[
T_{sr} = \begin{bmatrix}
\cos \delta & 0 & -\sin \delta \\
0 & 1 & 0 \\
\sin \delta & 0 & \cos \delta
\end{bmatrix},
\]  
\[\text{(2.92)}\]

and for the left hand side solar panel, the transformation matrix is (see Fig. 2.4):

\[
T_{sl} = \begin{bmatrix}
-\cos \delta & 0 & \sin \delta \\
0 & -1 & 0 \\
\sin \delta & 0 & \cos \delta
\end{bmatrix},
\]  
\[\text{(2.93)}\]

where \(\delta\) is the offset angle of the flexible solar panels (see Fig. 2.4).

These expressions are needed in Eqs. (2.61), (2.62) and (2.63).

Since the solar panels are flat plate structures, the transformation matrix for each panel will be the same for any plate element.

For application of the finite element method, the following choices are useful.

1. The solar panels will be divided into rectangular elements.
2. Each element has a uniform mass density \(\rho_i\).
3. Only out-of-plane deformations of the solar panels are considered.
4. External forces and torques on the solar panels are neglected. The rigid central body is allowed to rotate due to the external torques.
5. The \(y\) axis of the body fixed frame and the \(y\) axis of the solar panel local frame are parallel, or anti-parallel. We define that the \(x\) and \(y\) axes of the local reference frames are in the panel plane and the \(z\) axes are normal to the plane.
Fig. 2.4 A flexible spacecraft as used in simulation.

2.5.2 Element and nodal numbering algorithm

From the above definitions, the flexible solar panels are divided into rectangular elements. Each corner of an element is defined by a node and each element has as four nodes. For out-of-plane displacements we must consider for each node one translational displacement in the z direction and two rotational displacements about x and y directions. Therefore, for each element there are 12 degrees of freedom in total, see Fig. 2.6.

The shape of the deflected element is described by the displacement function matrix \( E_i \) that defines the deformations of the distributed points of the element as a function of the coordinates of these points (see Eq. (2.33)).

In the present case, the distributed displacements \( E_{po,p} \) have only one non-zero component, in the z direction of the local reference frame \( F_1 \) and the number of the degrees of freedom of nodes, see Fig. 2.6, is chosen to be 12.
It means that for element \( \mathbf{P}_{z,i} \) in Eq. (2.33) is an \( 1 \times 12 \) matrix and \( \mathbf{c} \) is a \( 12 \times 1 \) coefficient vector to be determined.

![Diagram of solar panel local frame definitions]

**Fig. 2.5. Solar panel local frame definitions.**

A often used choice of these displacement functions \( \mathbf{P}_{z,i} \) in Eq. (2.33), for out-of-plane-deformation problem is given in Ref. [12] by:

\[
\mathbf{P}_{z,i} = \begin{bmatrix}
1 & x & y & x^2 & xy & x^3 & x^2y & xy^2 & y^3 & x^3y & xy^3
\end{bmatrix}
\]

The coefficient vector \( \mathbf{c}_{z,i} \) in Eq. (2.33) can be written as:

\[
\mathbf{c}_{z,i} = \begin{bmatrix}
c_{i,1} & c_{i,2} & \ldots & c_{i,12}
\end{bmatrix}^T
\]

The matrix \( \mathbf{Z}_{z,i,k} \) in Eq. (2.34) can be obtained by using the displacement function matrix \( \mathbf{P}_{z,i} \) as:
\[
Z_{1,k} = \frac{1}{2} \begin{bmatrix}
2 \\
\partial / \partial y \\
-\partial / \partial x
\end{bmatrix} \cdot F_{1} \cdot (k \neq 1, 2, 3, 4)
\text{ at } x_{k}, y_{k}
\]

The matrix \( Z_{1} \) can be found by:

\[
Z_{1} = [Z_{1,1}^{T} Z_{1,2}^{T} Z_{1,3}^{T} Z_{1,4}^{T}]^{T}.
\]

Using (2.38), the shape function matrix \( \mathbf{C}_{i} \) is then obtained by using the expressions of \( F_{i} \) and \( Z_{i} \) as:

\[
\mathbf{C}_{i}^{T} =
\begin{bmatrix}
1 - \xi \eta - (3 - 2\xi) \xi^{2} (1 - \eta) + (1 - \xi) (3 - 2\eta) \eta^{2} \\
(1 - \xi) \eta (1 - \eta^{2}) \beta \\
-\xi (1 - \xi)^{2} (1 - \eta) \alpha \\
(1 - \xi) (3 - 2\eta) \eta^{2} + \xi (1 - \xi) (1 - 2\xi) \eta \\
- (1 - \xi) (1 - \eta) \eta^{2} \beta \\
-\xi (1 - \xi)^{2} \eta \alpha \\
(3 - 2\xi) \xi^{2} \eta - \xi \eta (1 - \eta) (1 - 2\eta) \\
-\xi (1 - \eta) \eta^{2} \beta \\
(1 - \xi) \xi^{2} \eta \alpha \\
(3 - 2\xi) \xi^{2} (1 - \eta) + \xi \eta (1 - \eta) (1 - 2\eta) \\
\xi \eta (1 - \eta)^{2} \beta \\
(1 - \xi) \xi^{2} (1 - \eta) \alpha
\end{bmatrix}, \quad (2.94)
\]

where \( 0 \leq \xi = \frac{x}{a} \leq 1 \) and \( 0 \leq \eta = \frac{y}{b} \leq 1 \).

Since the elements are all selected as the rectangular plate elements, the shape function matrix \( \mathbf{C}_{i} \) remain the same for every element.
Fig. 2.6 Plate element of the solar panel.

It should be noted, that other choices for $F_1$ are possible as well. The present choice has some drawbacks which will be discussed later, and therefore, also alternatives will be discussed.

In a plate flexure problem [12], the state of strain at any point may be represented by three components, namely the curvature in the $x$ direction, the curvature in the $y$ direction and the twist. The curvature in the $x$ direction is equal to the rate of change of the slope in the $x$ direction with respect to $x$ and is:

$$-\frac{3}{\delta x^2} \frac{\partial w}{\partial x^2} = -\frac{3}{\delta x^2} \frac{\partial w}{\partial x^2}, \quad (2.95)$$

where $w$ is the component of $R_{p0,p1}$ in the $z$ direction.

Similarly, the curvature in the $y$ direction is:

$$-\frac{\partial}{\partial y^2} \frac{\partial w}{\partial y^2} = -\frac{\partial^2 w}{\partial y^2}, \quad (2.96)$$
Finally, the twist is equal to the rate of change of the slope in the \( x \) direction with respect to \( y \) and is:

\[
\frac{\partial}{\partial y} \left( \frac{\partial w}{\partial x} \right) = -\frac{\partial^2 w}{\partial x \partial y} .
\]  

(2.97)

The state of 'strain' in the element can thus be represented as:

\[
\varepsilon = \begin{bmatrix}
\varepsilon_{xx} \\
\varepsilon_{yy} \\
\varepsilon_{xy}
\end{bmatrix} = \begin{bmatrix}
-\frac{\partial^2 w}{\partial x^2} \\
-\frac{\partial^2 w}{\partial y^2} \\
2 \frac{\partial^2 w}{\partial x \partial y}
\end{bmatrix} = B \varepsilon_1 w .
\]  

(2.98)

The internal torques \( T_x \) and \( T_y \) each act on two sides of the element. Similarly the twisting torques \( T_{xy} \) and \( T_{yx} \) each act on two sides (see, Fig. 2.7). But, since \( T_{xy} \) is equal to \( T_{yx} \), one of these twists, say \( T_{xy} \) can be considered to act on all four sides. This effect is allowed for, simply by doubling the twist term in the strain vector in (2.98). Moreover, \( w \) is just a component of the distributed displacement vector \( R_{p0,p,l} \) in the \( z \) direction and:

\[
w = C \varepsilon_1 \cdot d_1 ,
\]  

(2.99)

where \( d \) is the nodal displacement vector of the element, i.e.:

\[
d_1 = [d_{1,1} \ d_{1,2} \ \ldots \ d_{1,12}]^T .
\]  

(2.100)

The internal 'stresses' are directly related to the bending and twisting moments. Thus, for a plate flexure problem the state of 'stress' can be represented by the three moments \( T_x \), \( T_y \) and \( T_{xy} \):
\[
\sigma_i = \begin{bmatrix}
T_x \\
T_y \\
T_{xy}
\end{bmatrix}
\] (2.101)

where \(T_x\) and \(T_y\) are called internal bending moments and \(T_{xy}\) is called the internal twisting moment. If a small rectangular portion from within the finite element is considered, these internal bending and twisting moments are shown as in Fig. 2.7. \([T_{xy} = T_{yx}]\).

In this case, let \(D_x\) and \(D_y\) be the flexure rigidities in the \(x\) and \(y\) directions, respectively, \(D_1\) is a coupling rigidity representing a Poisson's ratio type of effect, and \(D_{xy}\) is the torsional rigidity. An isotropic plate, see Ref. [8], [9] and [10], has the same elastic properties in all directions and for this case:

\[
D_x = D_y = D = \frac{E t^3}{12 (1 - \nu^2)} ,
\] (2.102)

\[
D_1 = \nu D ,
\] (2.103)

and:

\[
D_{xy} = \frac{1}{2} (1 - \nu) D ,
\] (2.104)

where \(t\) is the thickness of the element, \(\nu\) is the Poisson's ratio and \(E\) is the Young's modulus.

The matrix \(R_1\) in Eq. (2.71) in this case is:

\[
R_1 = \begin{bmatrix}
D_x & D_1 & 0 \\
D_1 & D_y & 0 \\
0 & 0 & D_{xy}
\end{bmatrix}
\]
The matrix $G_{\dot{e}_i}Z_{\dot{e}_i}^{-1}$ now can be obtained as a 12 by 3 matrix as follows:

$$G_{\dot{e}_i}Z_{\dot{e}_i}^{-1} = \begin{bmatrix}
g_{1,1} & g_{1,2} & \cdots & g_{1,12} \\
g_{2,1} & g_{2,2} & \cdots & g_{2,12} \\
g_{3,1} & g_{3,2} & \cdots & g_{3,12}
\end{bmatrix},$$

(2.106)

The elements of $G_{\dot{e}_i}Z_{\dot{e}_i}^{-1}$ are listed in Tab. 2.1.

Substituting (2.105) and (2.106), the symmetric stiffness matrix of the discussed element in a case of plate flexure can be written as:

$$K_{\dot{e}_i} = \frac{E \, t^3}{12(1 - v^2)ab} \begin{bmatrix}
K_{II} & \text{symmetric} \\
K_{IJ} & K_{JJ}
\end{bmatrix}.$$  

(2.107)

The submatrices $K_{II}$, $K_{IJ}$ and $K_{JJ}$ are written out in Appendix B.
<table>
<thead>
<tr>
<th>i,j</th>
<th>$g_{i,j}$</th>
<th>i,j</th>
<th>$g_{i,j}$</th>
<th>i,j</th>
<th>$g_{i,j}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1,1</td>
<td>$(1 - 2ξ) (1 - η) \frac{6}{a^2}$</td>
<td>2,1</td>
<td>$(1 - ξ) (1 - 2η) \frac{6}{b^2}$</td>
<td>3,1</td>
<td>$\left[1 - 6ξ (1 - ξ) - 6η (1 - η) \frac{2}{ab}\right]$</td>
</tr>
<tr>
<td>1,2</td>
<td>0</td>
<td>2,2</td>
<td>$(1 - ξ) (2 - 3η) \frac{2}{b}$</td>
<td>3,2</td>
<td>$(1 - 4η + 3η^2) \frac{2}{a}$</td>
</tr>
<tr>
<td>1,3</td>
<td>$-(2 - 3ξ) (1 - η) \frac{2}{a}$</td>
<td>2,3</td>
<td>0</td>
<td>3,3</td>
<td>$-(1 - 4ξ + 3ξ^2) \frac{2}{b}$</td>
</tr>
<tr>
<td>1,4</td>
<td>$(1 - 2ξ) η \frac{6}{a^2}$</td>
<td>2,4</td>
<td>$-(1 - ξ) (1 - 2η) \frac{6}{b^2}$</td>
<td>3,4</td>
<td>$\left[-1 + 6ξ (1 - ξ) + 6η (1 - η) \frac{2}{ab}\right]$</td>
</tr>
<tr>
<td>1,5</td>
<td>0</td>
<td>2,5</td>
<td>$(1 - ξ) (1 - 3η) \frac{2}{b}$</td>
<td>3,5</td>
<td>$-η (2 - 3η) \frac{2}{a}$</td>
</tr>
<tr>
<td>1,6</td>
<td>$-(2 - 3ξ) η \frac{2}{a}$</td>
<td>2,6</td>
<td>0</td>
<td>3,6</td>
<td>$(1 - 4η + 3η^2) \frac{2}{b}$</td>
</tr>
<tr>
<td>1,7</td>
<td>$-(1 - 2ξ) η \frac{6}{a^2}$</td>
<td>2,7</td>
<td>$-ξ (1 - 2η) \frac{6}{b^2}$</td>
<td>3,7</td>
<td>$1 - 6ξ (1 - ξ) - 6η (1 - η) \frac{2}{ab}$</td>
</tr>
<tr>
<td>1,8</td>
<td>0</td>
<td>2,8</td>
<td>$ξ (1 - 3η) \frac{2}{b}$</td>
<td>3,8</td>
<td>$η (2 - 3η) \frac{2}{a}$</td>
</tr>
<tr>
<td>1,9</td>
<td>$-(1 - 3ξ) η \frac{2}{a}$</td>
<td>2,9</td>
<td>0</td>
<td>3,9</td>
<td>$-ξ (2 - 3ξ) \frac{2}{b}$</td>
</tr>
<tr>
<td>1,10</td>
<td>$-(1 - 2ξ) (1 - η) \frac{6}{a^2}$</td>
<td>2,10</td>
<td>$ξ (1 - 2η) \frac{6}{b^2}$</td>
<td>3,10</td>
<td>$\left[-1 + 6ξ (1 - ξ) + 6η (1 - η) \frac{2}{ab}\right]$</td>
</tr>
<tr>
<td>1,11</td>
<td>0</td>
<td>2,11</td>
<td>$ξ (2 - 3η) \frac{2}{b}$</td>
<td>3,11</td>
<td>$-(1 - 4η + 3η^2) \frac{2}{a}$</td>
</tr>
<tr>
<td>1,12</td>
<td>$-(1 - 3ξ) (1 - η) \frac{2}{a}$</td>
<td>2,12</td>
<td>0</td>
<td>3,12</td>
<td>$ξ (2 - 3ξ) \frac{2}{b}$</td>
</tr>
</tbody>
</table>

Tab. 2.1 Elements of matrix $G_{s1 s1}^{-1}$ in Eq. (2.106)

The shape function (2.94), however, has a drawback since the slopes normal to the edges of elements are discontinuous as shown in Fig. 2.8.
To overcome this problem, Bogner [13] and [14] introduced a 'compatible' element shape function matrix. In their work the shape function matrix can guarantee, that deflections and slopes are all continuous on the edges of elements. The shape function matrix they obtained is of the following form:

\[
\mathbf{C}_{\xi_1}^T = 
\begin{bmatrix}
(1 + 2\xi) (1 - \xi)^2 (1 + 2\eta) (1 - \eta)^2 \\
(1 + 2\xi) (1 - \xi)^2 \eta (1 - \eta)^2 b \\
- \xi (1 - \xi)^2 (1 + 2\eta) (1 - \eta)^2 a \\
(1 + 2\xi) (1 - \xi)^2 (3 - 2\eta) \eta^2 \\
- (1 + 2\xi) (1 - \xi)^2 (1 - \eta) \eta^2 b \\
- \xi (1 - \xi)^2 (3 - 2\eta) \eta^2 a \\
(3 - 2\xi) \xi^2 (3 - 2\eta) \eta^2 \\
- (3 - 2\xi) \xi^2 (1 - \eta) \eta^2 b \\
(1 - \xi) \xi^2 (3 - 2\eta) \eta^2 a \\
(3 - 2\xi) \xi^2 (1 + 2\eta) (1 - \eta)^2 \\
(3 - 2\xi) \xi^2 \eta (1 - \eta)^2 b \\
(1 - \xi) \xi^2 (1 + 2\eta) (1 - \eta)^2 a
\end{bmatrix}
\]

\[ (2.108) \]

The matrix \( G_{\xi_1} Z_{\xi_1}^{-1} \) can be obtained from (2.108) as follows:

\[
G_{\xi_1} Z_{\xi_1}^{-1} = 
\begin{bmatrix}
g_{1,1} & g_{1,3} & \cdots & g_{1,12} \\
g_{2,1} & g_{2,2} & \cdots & g_{2,12} \\
g_{3,1} & g_{3,2} & \cdots & g_{3,12}
\end{bmatrix}
\]

\[ (2.109) \]

The elements of matrix \( G_{\xi_1} Z_{\xi_1}^{-1} \) in (2.109) are given in Tab. 2.2.

The stiffness matrix of this type of shape function is also written out in Appendix A.

The mass matrices for two types of the shape function matrices (2.94) and (2.108) can be obtained from (2.42). The detailed expressions of the mass
matrices of the discussed element for both incompatible and compatible shape function matrices are given in Appendix B as well.

![Deflected elements with slopes normal to edge are discontinuous]

**Fig. 2.8 Discontinuing slopes normal to the edges.**

<table>
<thead>
<tr>
<th>i, j</th>
<th>( g_{i,j} )</th>
<th>i, j</th>
<th>( g_{i,j} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1,1</td>
<td>((1 - 2\xi)(1 + 2\eta)(1 - \eta)^2 \frac{6}{a^2})</td>
<td>1,2</td>
<td>((1 - 2\xi)(1 - \eta)^2 \frac{6b}{a^2})</td>
</tr>
<tr>
<td>1,3</td>
<td>(-(2 - 3\xi)(1 + 2\eta)(1 - \eta)^2 \frac{2}{a})</td>
<td>1,4</td>
<td>((1 - 2\xi)(3 - 2\eta) \eta^2 \frac{6}{a^2})</td>
</tr>
<tr>
<td>1,5</td>
<td>(-(1 - 2\xi)(1 - \eta) \eta^2 \frac{6b}{b^2})</td>
<td>1,6</td>
<td>(-(2 - 3\xi)(3 - 2\eta) \eta^2 \frac{2}{a})</td>
</tr>
<tr>
<td>1,7</td>
<td>(-(1 - 2\xi)(3 - 2\eta) \eta^2 \frac{6}{a^2})</td>
<td>1,8</td>
<td>((1 - 2\xi)(1 - \eta) \eta^2 \frac{6b}{a^2})</td>
</tr>
<tr>
<td>1,9</td>
<td>(-(1 - 3\xi)(3 - 2\eta) \eta^2 \frac{2}{a})</td>
<td>1,10</td>
<td>(-(1 - 2\xi)(1 + 2\eta)(1 - \eta)^2 \frac{6}{a^2})</td>
</tr>
<tr>
<td>1,11</td>
<td>((1 - 2\xi)(1 + 2\eta)(1 - \eta)^2 \frac{6b}{a})</td>
<td>1,12</td>
<td>(-(1 - 3\xi)(1 + 2\eta)(1 - \eta)^2 \frac{2}{a})</td>
</tr>
</tbody>
</table>
(continued)

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2,1</td>
<td>(1 + 2ξ)(1 - ξ)^2(1 - 2η) ( \frac{6}{b^2} )</td>
<td>2,2</td>
</tr>
<tr>
<td>2,3</td>
<td>-ξ(1 - ξ)^2(1 - 2η) ( \frac{6a}{b^2} )</td>
<td>2,4</td>
</tr>
<tr>
<td>2,5</td>
<td>(1 + 2ξ)(1 - ξ)^2(1 - 3η) ( \frac{2}{b} )</td>
<td>2,6</td>
</tr>
<tr>
<td>2,7</td>
<td>-(3 - 2ξ)ξ^2(1 - 2η) ( \frac{6}{b^2} )</td>
<td>2,8</td>
</tr>
<tr>
<td>2,9</td>
<td>-(1 - ξ)ξ^2(1 - 2η) ( \frac{6a}{b^2} )</td>
<td>2,10</td>
</tr>
<tr>
<td>2,11</td>
<td>(3 - 2ξ)ξ^2(2 - 3η) ( \frac{2}{b} )</td>
<td>2,12</td>
</tr>
<tr>
<td>3,1</td>
<td>-ξ(1 - ξ)n(1 - n) ( \frac{72}{ab} )</td>
<td>3,2</td>
</tr>
<tr>
<td>3,3</td>
<td>-(1 - ξ)(1 - 3ξ)n(1 - n) ( \frac{12}{b} )</td>
<td>3,4</td>
</tr>
<tr>
<td>3,5</td>
<td>-ξ(1 - ξ)n(2 - 3n) ( \frac{12}{a} )</td>
<td>3,6</td>
</tr>
<tr>
<td>3,7</td>
<td>-ξ(1 - ξ)n(1 - n) ( \frac{72}{ab} )</td>
<td>3,8</td>
</tr>
<tr>
<td>3,9</td>
<td>-ξ(2 - 3ξ)n(1 - n) ( \frac{12}{b} )</td>
<td>3,10</td>
</tr>
<tr>
<td>3,11</td>
<td>-ξ(1 - ξ)(1 - n)(1 - 3n) ( \frac{12}{a} )</td>
<td>3,12</td>
</tr>
</tbody>
</table>

Tab. 2.2 Elements of matrix \( \mathbf{G}_{i1} \mathbf{Z}^{-1}_{i1} \) in Eq. (2.109).

2.5.3 Assembled model

The mass and the stiffness matrices given in Appendix B are only valid for one particular element. To evaluate the assembled model of the flexible spacecraft dynamics, the total mass matrix \( \mathbf{M}_a \), stiffness matrix \( \mathbf{K}_a \), total damping matrix \( \mathbf{D}_a \), the flexible solar panel inertia matrix \( \mathbf{I}_{sa} \) and the coupling matrix \( \mathbf{A}_a \) must be obtained. In Section 2.1 the rigid body inertia matrix \( \mathbf{I}_{rb} \)
has been obtained. This matrix should retain its form with integrals, because the mass density of the rigid main body of a spacecraft cannot be assumed to be uniform due to e.g. discrete instruments, so that the components of this matrix will be treated as unknown parameters to be identified.

However, the inertia matrix of the flexible solar panels can be evaluated since the mass density of the solar panels may be assumed to be uniform (see, e.g. [7]).

The coupling matrix \( A \) shown in (2.62), can also be calculated because the shape function matrix \( C_{s1} \) is chosen (see (2.94) and (2.108)) and the integrals can be evaluated.

As discussed in Subsection 2.5.2, the distributed displacement vector \( R_{po,p,l} \) for plate flexure deformations has only one non-zero component, along the z direction in the local reference frame, i.e.:

\[
R_{po,p,l} = \begin{bmatrix} 0 \\ 0 \\ w \end{bmatrix} = w \bar{k}_l = \begin{bmatrix} 0 \\ 0 \\ C_{s1} \end{bmatrix} \cdot \bar{d}_l, \quad (2.110)
\]

Therefore, (2.46), (2.47) and (2.48) become:

\[
W_{s1} = \begin{bmatrix} 0 \\ 0 \\ \int \frac{C_{s1}}{E_1} \, dm \end{bmatrix}, \quad (2.111)
\]

\[
V_{s1} = \int_{E_1} R_{po,p,l} \cdot \begin{bmatrix} 0 \\ 0 \\ C_{s1} \end{bmatrix} \, dm, \quad \text{and:} \quad (2.112)
\]
where \( x_{oi} \) and \( y_{oi} \) are the components of \( R_{b,o,i} \) along the x axis and y axis in the local reference frame. It should be pointed out, that \( x_{oi} \) and \( y_{oi} \) vary with the element i. For different elements, \( x_{oi} \) and \( y_{oi} \) may be different since they depend on different locations of the local reference frames as shown in Fig. 2.9.

Using the local reference frame numbering system shown in Fig. 2.9, we can obtain:

\[
y_{oi} = 1 + \frac{i - q}{m_a} b \quad , \quad (q = 1, 2, \ldots, m_a) \quad , \quad (2.114)
\]

\[
x_{oi} = [i - (k - \frac{1}{2}) m_a - 1] a \quad , \quad (k = 1, 2, \ldots, m_b) \quad , \quad (2.115)
\]

where \( m_a \) and \( m_b \) are the number of elements along the x and y axes, \( l \) is the distance from the origin of the body fixed frame to the panel edge, \( a \) and \( b \) are the sizes of the element along the x and y axes respectively.

To give the exact expression of the coupling matrix \( A_{i1} \), the integrals in Eq. (2.46) and (2.47) have to be calculated.

Referring to the continuous shape function matrix (2.108), it follows that:

\[
\int_{E_1} C_{i1} \, dm = \frac{p_{tab}}{24} \begin{bmatrix} 6 & b & -a & 6 & -b & -a & 6 & -b & a \end{bmatrix}^T , \quad (2.116)
\]

\[
\int_{E_1} y C_{i1} \, dm = \frac{p_{tab}^2}{240} \begin{bmatrix} 18 & 4b & -6a & 42 & -6b & -7a & 42 & -6b & 7a & 18 & 4b & 3a \end{bmatrix}^T , \quad (2.117)
\]
\[
\sum_{i} c_i \text{dm} = \frac{\rho t a^2 b}{240} \begin{bmatrix} 18 & 3b & -3a & 18 & -3b & -4a \\
42 & -7b & 6a & 42 & 7b & 6a \end{bmatrix}^T.
\]

(2.118)

Therefore, the coupling matrix \( A_{\mathbf{1}} \) can be obtained as follows:

\[
A_{\mathbf{1}}^T = \frac{\rho t a b}{24} \begin{bmatrix}
6y_{o1} + \frac{9}{5} b & -6x_{o1} - \frac{3}{5} a & 0 \\
by_{o1} + \frac{2}{5} b & -bx_{o1} - \frac{3}{10} ab & 0 \\
-ay_{o1} - \frac{3}{5} ab & ax_{o1} + \frac{3}{10} a^2 & 0 \\
6y_{o1} + \frac{21}{5} b & -6x_{o1} - \frac{9}{5} a & 0 \\
-by_{o1} - \frac{3}{5} b^2 & bx_{o1} + \frac{3}{10} ab & 0 \\
-ay_{o1} - \frac{7}{10} ab & ax_{o1} + \frac{2}{5} a^2 & 0 \\
6y_{o1} + \frac{21}{5} b & -6x_{o1} - \frac{21}{5} a & 0 \\
-by_{o1} - \frac{3}{5} b^2 & bx_{o1} + \frac{7}{10} ab & 0 \\
ay_{o1} + \frac{7}{10} ab & -ax_{o1} - \frac{3}{5} a^2 & 0 \\
6y_{o1} + \frac{9}{5} b & -6x_{o1} - \frac{21}{5} a & 0 \\
by_{o1} + \frac{2}{5} b^2 & -bx_{o1} - \frac{7}{10} ab & 0 \\
ay_{o1} + \frac{3}{10} ab & -ax_{o1} - \frac{3}{5} a^2 & 0
\end{bmatrix}.
\]

(2.119)
The inertia matrix of (2.50) can be calculated as follows. Since vector $\mathbf{R}_{b,o,l}$ is in the x-y plane of the local reference frame under consideration, see Fig. (2.4), it follows that:

$$\mathbf{R}_{b,o,l}^T \cdot \mathbf{R}_{b,o,l} = \begin{bmatrix} y_{oi}^2 & -x_{oi}y_{oi} & 0 \\ -x_{oi}y_{oi} & x_{oi}^2 & 0 \\ 0 & 0 & y_{oi}^2 + x_{oi}^2 \end{bmatrix}, \quad (2.120)$$

where:

$$\mathbf{R}_{b,o,l} = \begin{bmatrix} 0 & 0 & y_{oi} \\ 0 & 0 & -x_{oi} \\ -y_{oi} & x_{oi} & 0 \end{bmatrix}. \quad (2.120a)$$
Using (2.120a), the symmetric matrix \( I_{E_1} \) in (2.52) can be obtained as:

\[
I_{E_1} = \rho_{abt} \begin{bmatrix}
I_{E_{11}} & \text{symmetric} \\
I_{E_{21}} & I_{E_{22}} \\
I_{E_{31}} & I_{E_{32}} & I_{E_{33}}
\end{bmatrix},
\]

(2.121)

where the elements of \( I_{E_1} \) are obtained as follows:

\[
I_{E_{11}} = \frac{1}{12} t^2 + \frac{1}{3} b^2 + y_{oi} b,
\]

\[
I_{E_{21}} = -\frac{1}{2} \left( \frac{1}{2} a b + y_{oi} a + x_{oi} b \right),
\]

\[
I_{E_{31}} = I_{E_{32}} = 0,
\]

\[
E_{22} = \frac{1}{12} t^2 + \frac{1}{3} a^2 + x_{oi} a,
\]

\[
I_{E_{33}} = \frac{b^2}{3} + \frac{a^2}{3} + y_{oi} b + x_{oi} a.
\]

Substituting (2.120) and (2.121) in (2.51), the symmetric inertia matrix of element \( i \) is written as:

\[
I_i = \rho_{abt} \begin{bmatrix}
I_{11} & \text{symmetric} \\
I_{11} & I_{22} \\
I_{31} & I_{32} & I_{33}
\end{bmatrix},
\]

(2.122)

where the elements of \( I_i \) are obtained as follows:

Therefore, there are ten parameters to be identified in the rigid main body of the spacecraft.

2.6.2 Parameter to be identified in the flexible substructures
\[ I_{11} = y_{o1}^2 + \frac{1}{12} t^2 + \frac{1}{3} b^2 + y_{o1} b \ , \]
\[ I_{21} = -\frac{1}{2} \left( \frac{1}{2} a b + y_{o1} a + x_{o1} b \right) - x_{o1} y_{o1} \ , \]
\[ I_{31} = I_{32} = 0 \ , \]
\[ I_{22} = x_{o1}^2 + \frac{1}{12} t^2 + \frac{1}{3} a^2 + x_{o1} a \ , \]
\[ I_{33} = y_{o1}^2 + x_{o1}^2 + \frac{1}{3} b^2 + \frac{1}{3} a^2 + y_{o1} b + x_{o1} a \ . \]

Substituting the matrices (2.122), (2.92) and (2.93) in (2.61) and (2.62) and referring to the Appendix A to give the expressions of the mass and

different parameters, the total dynamical model takes the form

82

There are six parameters to be identified.
The elements might be different from each other in the most general case, therefore, the parameters to be identified in the flexible substructures are:

\[ P_a = \begin{bmatrix} P_1^T & P_2^T & P_3^T & \ldots & P_N^T \end{bmatrix}^T \ . \] (2.125)

If the flexible substructures contain N elements, in total, there are 6 x N parameters to be identified.
However, this most general case is not common. In practice, a large number of elements will be treated as identical elements, due to the same material of the flexible substructure being involved. In this case the number of parameters to be identified will be greatly reduced. For example, if we select the elements of the solar panels of the spacecraft discussed in the last section to be identical, the only parameters to be identified will be:

\[ P_a = [b \ a \ t \ E \ \nu \ \rho \ \alpha \ \beta]^T \ . \] (2.126)

This means that the parameters to be identified in the whole flexible substructure are reduced to the parameters of a single element.

2.6.3 Parameters to identified in the total dynamical model

The parameters to be identified in the dynamical model of the entire
\[ b a t E \nu \rho \alpha \beta \]^{T}. \quad (2.128)

These parameters are only the structural parameters in the flexible spacecraft dynamical model described in Eq. (2.91). In the next chapter we shall discuss additional parameters in the model to be used for parameter identification.

The number of the parameters to be identified discussed in this section shows that for each element the number of the parameters to be identified can be reduced to six (see Eq. (2.124)). These parameters remain the same even when the element is a three dimensional element since the parameters of a three dimensional element are also the dimensions of the element \((a, b, t)\), the mass density \(\rho\) and material rigidities \(E\) and \(\nu\).

To show the reduction of number of the parameters to be identified, we examine the flexible spacecraft with two symmetrical solar panels as discussed in the last section. If the panels are divided into 32 elements, the number of displacements for the two solar panels is 162, since each node has three displacements as discussed in the last section. It means that the mass, damping and stiffness matrices are all 162 by 162 symmetric matrices. If the finite element method is not used to calculate the mass matrix \(M\) and the stiffness matrix \(K\), all the upper triangular elements of the two matrices have to be estimated. The parameters to be identified in these two matrices are then 26,406 since there is no any a priori knowledge about these matrices. If the finite element method is used to analyse the structure dynamics of these two solar panels, each element has six parameters to be identified, see Eq. (2.124). The number of the parameters in the mass matrix \(M\) and the stiffness matrix \(K\) becomes 192 when all the selected elements are different. In practice, it is common that large number of the elements are selected to be identical due to the structural type and the same material. For example, if all the elements are chosen to be identical for the two solar panels, the parameters to be identified in the two matrices, \(M\) and \(K\) are only six.

The canonical form of these matrices discussed in [16] and [17] is based on the so-called modal analysis which discusses the model natural frequencies, damping ratios and modal displacements. When the matrices \(M\) and \(K\) are 162 by
162 matrices, there are 162 natural frequencies and 162 damping ratios need to be estimated. Then 324 parameters are selected as the unknown parameters for the canonical form.

Using the finite element analysis, some time the parameters selected in the flexible substructures do not change much after launching [18]. For example, the dimensions of the structures and the material properties (mass density, Young's modulus and Poisson ratio) can be measured before launching and they do not change much in space. In this case the parameters to be identified in the flexible spacecraft dynamical model are only the parameters of the rigid main body and the damping coefficients if necessary. The mass and stiffness matrices can be calculated by using the finite element method. Therefore, the parameters to be identified can be reduced to a reasonable value for the parameter estimation procedures (see Chapter 3) to handle.

References


2.5 Van Woerkom, P.Th.L.M. "Hybrid Equations for Translational and Rotational Motion of Flexible Spacecraft", NLR TR 81116 U National Aerospace Laboratory (NLR), The Netherlands.


2.18 Van Den Broek, P. Ph. Private Communications.


CHAPTER 3

MAXIMUM LIKELIHOOD ESTIMATION OF FLEXIBLE SPACECRAFT PARAMETERS

3.1 Introduction

In Chapter 2, a general dynamical model of flexible spacecraft was obtained. This model can be applied to a large number of flexible spacecraft, since all possible motions of the rigid main body and the flexible substructures are considered (six degrees of freedom for the rigid main body and each node considered in the flexible substructures). This model can also be applied to space structures with different kinds of flexible appendages, such as the European Space Station Columbus (see Fig. 3.2), the Earth-Observation Platform (see Fig. 3.3), etc., since these space structures can also be considered as combinations of rigid bodies and flexible substructures. Therefore, the model discussed in the previous chapter is also suitable for these space structures. As shown in Chapter 2, the number of parameters to be identified is still manageable when a large number of so-called finite elements may be defined to be identical or calculate the system matrices (mass, stiffness and coupling) directly from the finite element method with the known parameters from ground tests.

The remaining question of Chapter 2 is whether this model is applicable to the design of spacecraft state estimators or system control systems. In general, the dynamical model of flexible spacecraft in ordinary differential equation form, as generated with the finite element method, will be of very high order. The underlying reason is, that in order to guarantee a 'sufficiently' accurate mathematical model, a large number of elements must be selected. This reduces the chance of overlooking a certain flexible mode which may be of importance, e.g., for the actual operation of the control system.

As an example, a spacecraft with two symmetrical solar panels as discussed in Chapter 2 is considered. If it is assumed that each panel is of dimension $4 \times 48 \text{ m}^2$, and each solar panel is divided into rectangular elements of dimension $0.5 \times 1 \text{ m}^2$ each to obtain the 'sufficiently' accurate model of the flexible spacecraft dynamics, the two solar panels will contain 768 elements. As shown in Chapter 2, each node has three degrees of freedom (one
translational displacement along the z axis of the flexible substructure local reference frame and two rotational displacements about the x and y axes of the same frame) and, therefore, the number of first order differential equations will be 5292, excluding the equations for the rigid central body of the spacecraft. Obviously, design and implementation of a control system for a system model of such high order may be quite difficult or even impossible, since the control algorithms are limited by computer wordlength, memory size and speed and by the accuracy of iterative procedures. These limitations result in maximum solvable dimensions of Riccati or Lyapunov equations. The 'dimensionality problem' becomes even more severe, if the number of finite elements must be increased due to the possible larger dimensions of future spacecraft. This means, that the order of the dynamical model of a flexible spacecraft based on finite element analysis must be reduced to an acceptable order in practical applications (see e.g. [25], [36], [37] and [38]).

Unfortunately, when the order of the mathematical model is reduced, modelling errors become unavoidable. Furthermore, these errors may be unknown, therefore, the reduced order model will exhibit certain so-called model uncertainties. If these model uncertainties are not taken into account, a satisfactory fit of the reduced model response to the actual spacecraft response may, in general, not be obtained. Furthermore, model errors may result in biases of the estimated model parameters, see e.g. [14]. To take account of the model uncertainties resulting from order reduction, a certain mathematical representation, describing the static or dynamic behaviour of the modelling errors, should be included in the reduced order model used for parameter identification. In general, the parameters in the mathematical representation of the modelling errors must be assumed to be unknown. Therefore, these parameters must be estimated simultaneously with the unknown parameters in the reduced order model.

As mentioned in Chapter 1, this study focuses on the maximum likelihood parameter identification of flexible spacecraft. The maximum likelihood parameter estimator (see e.g. [12], [16], [17] and [19]), however, should be extended to the case of a system with correlated process and measurement noises. As shown in Subsection 3.2.3, the correlation between process and measurement noise is a consequence of the introduction of acceleration or rather, specific force measurements on the flexible substructures.
This chapter is organized as follows.

Section 3.2 discusses the model to be used for maximum likelihood (ML) identification. In section 3.3, several 'versions' of ML algorithms for parameter estimation are obtained. To optimize the likelihood function, some optimization methods are discussed in Section 3.4, especially the Gauss-Newton Method. Section 3.5 discusses the identifiability of the flexible spacecraft model parameters and Section 3.6 gives the properties of the ML parameter estimator as described in the existing literature. Finally, detailed calculations of the sensitivity matrices of the model with respect to the parameters to be identified are presented in Section 3.7.

3.2 The Mathematical Model to be Used for ML Parameter Identification

A mathematical model of a flexible spacecraft, not only contains a set of dynamic equations of motion, which describe the dynamical behaviour of the spacecraft in space (see Chapter 2), but must also include a model of measurements and measurement errors. In Subsection 3.2.1, the possible instrumentation systems to be used for measuring flexible spacecraft motions are discussed.

Subsection 3.2.2 starts with the order reduction of the original high order dynamical model and then describes the actual lower order system to be used for parameter identification.

When the order of a system is reduced, modelling errors will be introduced in the system model. A suitable model of these modelling errors is discussed in Subsection 3.2.3. In practice, the model to be used for parameter identification is a discrete time system model, since the model inputs and the outputs are usually sampled at discrete time instants. The discrete time system model is derived in Subsection 3.2.4 from the continuous time system model in Subsection 3.2.3. Finally, the parameters to be estimated in the model are discussed in Subsection 3.2.4.

3.2.1 Instrumentation system and output equations of flexible spacecraft

The measurements of the motions of a flexible spacecraft in orbit can be divided into two parts, i.e.:

1. measurements of the motions of the rigid main body, and
2. measurements of the motions of the flexible substructures.
The motions of the rigid main body of a spacecraft can be measured using one
or more of the following transducers or techniques (see [7]):
1) gyros,
2) star telescopes and trackers, using star field measurements 'star
mapping',
3) sun sensors,
4) horizon sensors,
5) magnetometers,
6) gravity gradient sensors,
7) landmark telescopes, using landmark field measurements 'mapping', and
8) accelerometers.
Not all the instuments described above can be applied to the measurement of
the motions of the flexible substructures. The reason is, that most of the
transducers mentioned above are relatively heavy and can not readily be in-
stalled on flexible appendages. Only accelerometers can be light and small
enough to be considered as suitable transducers for the measurement of the
motions of flexible substructures in space for the time bing.
Considering the above instruments and measuring techniques, the observation
equations of the system model can be written as follows:

\[ \mathbf{y} = \mathbf{C}_1^* \mathbf{x} + \mathbf{C}_2^* \mathbf{\dot{x}} + \mathbf{y}_1, \quad (3.1) \]

where \( \mathbf{y} \) is a measurement vector which contains all possible transducer out-
puts, \( \mathbf{C}_1^* \) is an observation matrix related to acceleration or specific force
measurements (since \( \mathbf{x} \) includes accelerations or specific force of the
spacecraft, see Chapter 2), \( \mathbf{C}_2^* \) is an observation matrix related to measure-
ments of velocities and displacements of the spacecraft (see also Chapter 2)
and \( \mathbf{y}_1 \) is the corresponding measurement noise vector.
The measurement noise \( \mathbf{y}_1 \) in this study is assumed to be a white noise
process with zero mean and Gaussian distribution. If the actual measurement
noise would be a nonwhite noise, it could be represented as filtered white
noise, resulting in some additional filter states to be included in the system state vector (see [7] and [44]). Constant biases of the measurement transducers can be modelled as constant states which can be included in the system state vector as well (see [7], [42] and [44]). Therefore, (3.1) may be considered as a fairly general set of observation equations.

3.2.3 Order reduction of the flexible spacecraft dynamical model

If the mathematical description of flexible spacecraft dynamics, as generated from the finite element method discussed in Chapter 2, has to be as accurate as possible, the elements in the flexible substructures must be chosen relatively small. Consequently, a system model of very high order is obtained (see also Section 3.1). As mentioned in the introduction to this chapter, it may be necessary to reduce the order of such system models to a lower and more practical value.

Ref. [1] gives a comprehensive survey. It contains about 150 references on mathematical model order reduction procedures and describes five basic approaches in some detail.

A model order reduction algorithm related to parameter identification, as mentioned in [1], is the so-called Parameter Optimization approach. In this approach, a model with reduced order will be used to estimate the model parameters from measurement data of an actual spacecraft. The parameters to be estimated in the reduced order model are different from the parameters in the originally non-reduced model.

Problems that often arise using this approach are (see [1]):
1) convergence of the parameter optimization algorithms is not assured,
2) convergence to the global optimum of the criterion in parameter space is not assured, and
3) parameters may converge to wrong values.

These problems are caused by the unmodelled modes of the actual system being left out of consideration in the reduced order model, resulting in modelling errors. The optimization algorithms work well, only if the errors between the model and the actual spacecraft dynamics are zero or very small (see e.g. [28]). In the present work, the model errors resulting from model order reduction will explicitly be taken into account, see Section 3.1.
As mentioned above, the parameter to be estimated in the reduced order system model are different from the parameters in the originally non-reduced system model, as obtained in Eq. (2.91) in Chapter 2. The reason is, that the order of the system model is reduced and the parameters in Eq. (2.91) are re-determined. The parameters in the originally non-reduced order model have real physical meanings as described in the last chapter, and the number of the parameters to be estimated can be managed in an acceptable value, since the finite element method has been used for generating the system model, large number of parameters in the so-called mass, damping, stiffness and coupling matrices, see Chapter 2, can directly be calculated from the finite element method and only reasonable number of parameters are treated to be unknown parameters to be estimated. If the parameters to be estimated in the reduced order system model are expected to be the same physical meanings as the parameters in the non-reduced order system model, as described by Eq. (2.91), and as the same number of parameters as Eq. (2.91), the so-called 'Static Condensation' can be used to reduce the order of the system model, see e.g. [2], [3], [4] and [5]. The parameters in the reduced order model resulting from this approach are still the same as those in the high order system. This approach is discussed as follows.

Let Eq. (2.84) represent a dynamical model of high order:

\[
\begin{align*}
\mathbf{w}^T \ddot{\mathbf{r}} + \mathbf{a}^T (\dddot{\mathbf{w}} + \omega_0 \dot{\mathbf{w}}) + \mathbf{M} \dddot{\mathbf{d}} + \mathbf{D} \dot{\mathbf{d}} + \mathbf{K} \mathbf{d} &= \mathbf{F}_a,
\end{align*}
\]

where \(\dddot{\mathbf{w}}\) indicates the second derivative with respect to time and \(\mathbf{r} = \mathbf{r}_{i,b}\), see Eq. (2.84).

Let the displacement or deflection vector \(\mathbf{d}\) in (3.2) be divided into two parts:

\[
\begin{align*}
\mathbf{d} &= \begin{bmatrix}
\mathbf{d}_m \\
\mathbf{d}_s
\end{bmatrix}
\end{align*}
\]

where \(\mathbf{T}_{se}\) is a constant square matrix. Now it is assumed that the displacements \(\mathbf{d}_s\) depend in some unique way on the displacements \(\mathbf{d}_m\). The latter we
shall call therefore, \( \mathbf{d}_m \) the vector of 'master' displacements and \( \mathbf{d}_s \) the vector of 'slave' displacements, see [4]. If it is assumed that:

\[
\mathbf{d}_s = T^* \mathbf{d}_m ,
\]

(3.4)

\( T^* \) denoting a constant matrix of appropriate dimensions, (3.3) follows that:

\[
\mathbf{d} = T^* \mathbf{e} = \begin{bmatrix} \mathbf{U} \\ \mathbf{0} \end{bmatrix} \mathbf{d}_m = T^* \mathbf{e} \mathbf{d}_m.
\]

(3.5)

Premultiplying (3.2) by \( T^* \mathbf{e} \), after substitution of (3.5), leads to the so-called condensed system model:

\[
\begin{align*}
\mathbf{d}^T \mathbf{e} \mathbf{w}^T \mathbf{r} + \mathbf{d}^T \mathbf{e} \mathbf{A}^T (\ddot{\mathbf{e}} + \omega_0 \dot{\mathbf{e}}) + \mathbf{d}^T \mathbf{e} \mathbf{M} \mathbf{d}_m + \mathbf{d}^T \mathbf{e} \mathbf{D} \mathbf{d}_m + \\
\mathbf{d}^T \mathbf{e} \mathbf{K} \mathbf{d}_m = \mathbf{w}^T \mathbf{r} + \mathbf{A}^T (\ddot{\mathbf{e}} + \omega_0 \dot{\mathbf{e}}) + \mathbf{M} \mathbf{d}_m + \mathbf{D} \mathbf{d}_m + \mathbf{K} \mathbf{d}_m = \mathbf{F}^c,
\end{align*}
\]

(3.6)

in which the so-called condensed system matrices are:

\[
\mathbf{w}^T = T^* \mathbf{e} \mathbf{w}^T = T^* \begin{bmatrix} \mathbf{w}_m^T \\ \mathbf{w}_s^T \end{bmatrix} T,
\]

(3.7)

\[
\mathbf{A}^T = T^* \mathbf{e} \mathbf{A}^T = T^* \begin{bmatrix} \mathbf{A}_m^T \\ \mathbf{A}_s^T \end{bmatrix} T,
\]

(3.8)

\[
\mathbf{M}^c = T^* \mathbf{e} \mathbf{M} = T^* \begin{bmatrix} \mathbf{M}_{mm} \\ \mathbf{M}_{ms} \\ \mathbf{M}_{sm} \\ \mathbf{M}_{ss} \end{bmatrix} T,
\]

(3.9)
\[
D_c = T^T T_e D^T T_e T = T^T \begin{bmatrix}
D_{mm} & D_{ms} \\
D_{ms}^T & D_{ss}
\end{bmatrix} T,
\]
\[
K_c = T^T T_e K^T T_e T = T^T \begin{bmatrix}
K_{mm} & K_{ms} \\
K_{ms}^T & K_{ss}
\end{bmatrix} T,
\]
\[
F_c = T^T T_e F_a = T^T \begin{bmatrix}
F_m \\
F_s
\end{bmatrix},
\]

and the condensed external force vector is:

where \( F_m \) and \( F_s \) are the external force vectors, respectively acting on the master and slave degrees of freedom.

The important question is now, how to determine the relations between the 'slave' \( d_s \) and 'master' \( d_m \) deflections. A suitable assumption (see [4]) which can be justified by engineering intuition, is that the general pattern of deformation will follow that which would be obtained by imposing the external forces \( F_m \), pertaining to the "master" degrees of freedom, on an otherwise unloaded structure in static condition. Thus partitioning with \( r = \]
\[
\theta = \theta_0 = 0, \quad d = d_0 = 0_m \quad \text{and} \quad F_s = 0_m, \quad \text{Eq. (3.2)} \] can be written as:

\[
K = \begin{bmatrix}
K_{mm} & K_{ms} \\
K_{ms}^T & K_{ss}
\end{bmatrix}
\begin{bmatrix}
d_m \\
\begin{bmatrix}
d_s \\
0_m
\end{bmatrix}
\end{bmatrix} = \begin{bmatrix}
F_m \\
0_m
\end{bmatrix},
\]

From (3.12) it follows that:

\[
K_{ms}^T d_m + K_{ss} d_s = 0_m.
\]
so that the displacements $\mathbf{d}_s$ can be represented by:

$$
\mathbf{d}_s = -K_s^{-1}K_s^T\mathbf{d}_m.
$$  \hspace{1cm} (3.14)

Referring to (3.4) and assuming $K_s$ is of full rank, the transformation matrix $T_s^*$ can now be written as:

$$
T_s^* = -K_s^{-1}K_s^T.
$$  \hspace{1cm} (3.15)

An important characteristic of the 'condensed system model' (3.6) is that the condensed system matrices as defined above are closely related to the system matrices of the original, non-reduced model. This means that the parameters in the condensed system matrices still have a clear, physical interpretation in terms of the parameters in the mass, damping and stiffness matrices $M_s$, $D_s$ and $K_s$ respectively, as defined in Chapter 2.

The following consideration may guide us in the choice of master displacements. Master displacements should be selected such that, when the structure is loaded by a unit force, a so-called 'basic' displacement mode result. Our intent is, that a superposition of a number of basic modes will adequately describe actual displacements during dynamical responses in certain areas of engineering interest. For example, in control system design, the lower natural frequencies of the flexible structure may be of particular importance for control algorithm synthesis (see [1]). The method for finding the lower modes of the flexible structure is described in literature, see e.g. [4].

The selected master displacements should be easily measured for parameter estimation and the number of master displacements may be as low as 1/10 or 1/20 of the total number of displacements [2].
Plate without elimination.
Number of displacements is 90
(one translation and two
rotations at each node).

<table>
<thead>
<tr>
<th>mode</th>
<th>$\omega \sqrt{D/\rho \tau a}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.469</td>
</tr>
<tr>
<td>2</td>
<td>8.535</td>
</tr>
<tr>
<td>3</td>
<td>21.450</td>
</tr>
<tr>
<td>4</td>
<td>27.059</td>
</tr>
</tbody>
</table>

Nodes not ringed are eliminated.
Number of master displacements is 54.

<table>
<thead>
<tr>
<th>mode</th>
<th>$\omega \sqrt{D/\rho \tau a}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.470</td>
</tr>
<tr>
<td>2</td>
<td>8.540</td>
</tr>
<tr>
<td>3</td>
<td>21.559</td>
</tr>
<tr>
<td>4</td>
<td>27.215</td>
</tr>
</tbody>
</table>

All displacements eliminated,
except for translational dis-
placements at ringed nodes.
Number of master displacements is 18.

<table>
<thead>
<tr>
<th>mode</th>
<th>$\omega \sqrt{D/\rho \tau a}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.470</td>
</tr>
<tr>
<td>2</td>
<td>8.543</td>
</tr>
<tr>
<td>3</td>
<td>21.645</td>
</tr>
<tr>
<td>4</td>
<td>27.296</td>
</tr>
</tbody>
</table>

All displacements eliminated,
except for translational dis-
placements at ringed nodes.
Number of master displacements is 6.

<table>
<thead>
<tr>
<th>mode</th>
<th>$\omega \sqrt{D/\rho \tau a}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.473</td>
</tr>
<tr>
<td>2</td>
<td>8.604</td>
</tr>
<tr>
<td>3</td>
<td>22.690</td>
</tr>
<tr>
<td>4</td>
<td>29.490</td>
</tr>
</tbody>
</table>

Fig. 3.1. First four natural frequencies of a square cantilever plate with $D/\rho \tau a = 1$, see [4].
Static condensation is also used in the finite element method, to reduce the order of the dynamic system model in the modal analysis, as an eigenvalue economizer for searching the lower frequency modes of the dynamical system (see, e.g., [2], [3], [4] and [5]). Fig. 3.1 gives an example of the accuracy provided by the economizer. Triangular element and consistent mass matrices were used in this example, see [4].

The reduced order dynamical model of a flexible spacecraft can now be written as, see also Eqs. (2.91) and (3.6):

\[
F_r x_r = A_r x_r + B_r u_r ,
\]

\(\text{(3.16)}\)

where:

\[
F_r = \begin{bmatrix}
    Q & W_c & 0_{3 \times (6+m_r)} \\
    Q_T & I & A_c & 0_{3 \times (6+m_r)} \\
    W_c^T & A_c^T & M_c & 0_{m_r \times (6+m_r)} \\
    0_{(6+m_r) \times 3} & 0_{(6+m_r) \times 3} & 0_{(6+m_r) \times m_r} & U_{6+m_r}
\end{bmatrix}
\]

\[
A_r = \begin{bmatrix}
    0_{3} & -Q \omega & 0_{3 \times m_r} & 0_{3} & 0_{3} & 0_{3 \times m_r} \\
    0_{3} & -I \omega & 0_{3 \times m_r} & 0_{3} & 0_{3} & 0_{3 \times m_r} \\
    0_{m_r \times 3} & -A_c \omega & -D_c & 0_{m_r \times 3} & 0_{m_r \times 3} & -K_c \\
    U_{3} & 0_{3} & 0_{3 \times m_r} & 0_{3} & 0_{3} & 0_{3 \times m_r} \\
    0_{3} & U_{3} & 0_{3 \times m_r} & 0_{3} & 0_{3} & 0_{3 \times m_r} \\
    0_{m_r \times 3} & 0_{m_r \times 3} & U_{m_r} & 0_{m_r \times 3} & 0_{m_r \times 3} & 0_{m_r}
\end{bmatrix}
\]
\[
\begin{align*}
B_r &= \begin{bmatrix} [\mathbf{u} + \mathbf{m}]_r & \mathbf{0}_{(6+m)^T} \end{bmatrix}, \\
\chi_r &= \begin{bmatrix} \mathbf{e}^T_m & \mathbf{e}^T & \mathbf{d}^T_m & \mathbf{e}^T & \mathbf{d}^T_m \end{bmatrix}^T, \\
\upsilon_r &= \begin{bmatrix} \mathbf{e}^T_b & \mathbf{e}^T_b & \mathbf{e}^T_c \end{bmatrix}^T,
\end{align*}
\]

and \( m_r \) is the number of selected 'master' displacements.

The submatrices in \( B_r \) and \( A_r \) are obtained from (3.7) through (3.11) as follows:

\[
\begin{align*}
\mathbf{W}_c &= \mathbf{W}_m^T - \mathbf{K}_{ms} \mathbf{K}_{ss}^{-1} \mathbf{W}_s^T, \\
\mathbf{A}_c &= \mathbf{A}_m^T - \mathbf{K}_{ms} \mathbf{K}_{ss}^{-1} \mathbf{A}_s^T, \\
\mathbf{M}_c &= \mathbf{M}_{mm} - \mathbf{K}_{ms} \mathbf{K}_{ss}^{-1} \mathbf{M}_{ms} - \mathbf{M}_{ms} \mathbf{K}_{ss}^{-1} \mathbf{M}_{ss} + \mathbf{K}_{ms} \mathbf{K}_{ss}^{-1} \mathbf{M}_{ss} \mathbf{K}_{ss}^{-1} \mathbf{K}_{ms}^T, \\
\mathbf{D}_c &= \mathbf{D}_{mm} - \mathbf{K}_{ms} \mathbf{K}_{ss}^{-1} \mathbf{D}_{ms} - \mathbf{D}_{ms} \mathbf{K}_{ss}^{-1} \mathbf{D}_{ss} + \mathbf{K}_{ms} \mathbf{K}_{ss}^{-1} \mathbf{D}_{ss} \mathbf{K}_{ss}^{-1} \mathbf{K}_{ms}^T, \\
\mathbf{K}_c &= \mathbf{K}_{mm} - \mathbf{K}_{ms} \mathbf{K}_{ss}^{-1} \mathbf{K}_{ms}, \text{ and:} \\
\mathbf{F}_c &= \mathbf{F}_m.
\end{align*}
\]

Now we are turn to a basic problem in the parameter optimization approach. As mentioned in the beginning of this subsection, when a reduced order model is used, convergence of the parameter optimization algorithms to the global optimum is not assured, due to the fact that modelling errors resulting from order reduction are neglected all together.

If the finite dimensional model is obtained from the finite element method (see Chapter 2), modelling errors are unavoidable, even if high dimensional
models are employed, since from a theoretical point of view, the number of nodes should be infinite for a structure which is continuous. The errors may be negligible if the selected elements are sufficiently small and a correspondingly high order model is employed. However, if a reduced order model is used, modelling errors can no longer be neglected. They must be taken into account in the reduced order model.

Order reduction as generated by static condensation does not lead to model errors, only in 'static' conditions, i.e. when \( \ddot{r} = \dot{\theta} = \dot{\theta} = 0 \), \( \ddot{d} = \dot{d} = 0 \) and \( F_S = 0 \). Then the selected slave displacements do not contribute to the dynamic behaviour of the system. This is not the case in dynamic conditions. In dynamic conditions the slave displacements \( d_s \) can only be obtained by partitioning Eqs. (2.82), (2.83) in Chapter 2 and (3.2) as:

\[
\begin{bmatrix}
\ddots & \ddots & \ddots \\
\vdots & \ddots & \ddots \\
\ddots & \ddots & \ddots \\
\end{bmatrix}
\begin{bmatrix}
\ddots \\
\vdots \\
\ddots \\
\end{bmatrix}
= F_b, \quad (3.17)
\]

\[
\begin{bmatrix}
\ddots & \ddots & \ddots \\
\vdots & \ddots & \ddots \\
\ddots & \ddots & \ddots \\
\end{bmatrix}
\begin{bmatrix}
\ddots \\
\vdots \\
\ddots \\
\end{bmatrix}
= T_b, \quad (3.18)
\]

\[
\begin{bmatrix}
W^T_m \\
\vdots \\
W^T_S
\end{bmatrix}
\begin{bmatrix}
A^T_m \\
\vdots \\
A^T_S
\end{bmatrix}
\begin{bmatrix}
\ddots \\
\vdots \\
\ddots \\
\end{bmatrix}
\begin{bmatrix}
\ddots \\
\vdots \\
\ddots \\
\end{bmatrix}
+ \begin{bmatrix}
D_{mm} & D_{ms} \\
\vdots & \vdots \\
D_{ms} & D_{ss}
\end{bmatrix}
\begin{bmatrix}
\ddots \\
\vdots \\
\ddots \\
\end{bmatrix}
+ \begin{bmatrix}
K_{mm} & K_{ms} \\
\vdots & \vdots \\
K_{ms} & K_{ss}
\end{bmatrix}
\begin{bmatrix}
\ddots \\
\vdots \\
\ddots \\
\end{bmatrix}
= \begin{bmatrix}
F_m \\
\vdots \\
F_s
\end{bmatrix}, \quad (3.19)
\]
or:

\( \begin{align*}
\ddot{m} \dddot{r} + Q (\dddot{\theta} + \omega_0 \dot{\theta}) + W_m \dddot{d}_m + W_s \dddot{d}_s &= F_b, \\
Q^T \dddot{r} + \dddot{I} (\dddot{\theta} + \omega_0 \dot{\theta}) + \dddot{A}_m \dddot{d}_m + \dddot{A}_s \dddot{d}_s &= T_b, \\
W_m \dddot{r} + A_m^T (\dddot{\theta} + \omega_0 \dot{\theta}) + M_{mm} \dddot{d}_m + M_{ms} \dddot{d}_s + D_{mm} \dddot{d}_m + D_{ms} \dddot{d}_s + \\
+ K_{mm} \dddot{d}_m + K_{ms} \dddot{d}_s &= F_m, \\
W_s \dddot{r} + A_s^T (\dddot{\theta} + \omega_0 \dot{\theta}) + M_{ms}^T \dddot{d}_m + M_{ss} \dddot{d}_s + D_{ms}^T \dddot{d}_m + D_{ss} \dddot{d}_s + \\
+ K_{ms}^T \dddot{d}_m + K_{ss} \dddot{d}_s &= F_s.
\end{align*} \)

From (3.23) the following expressions for \( \dddot{d}_s, \dot{d}_s \) and \( \dddot{d}_s \) may be derived:

\( \begin{align*}
\dddot{d}_s &= F_s - K_{ss}^{-1} \left[ W_s \dddot{r} + A_s^T (\dddot{\theta} + \omega_0 \dot{\theta}) + M_{ms}^T \dddot{d}_m + M_{ss} \dddot{d}_s + \\
+ D_{ms}^T \dddot{d}_m + D_{ss} \dddot{d}_s \right] \), \\
\dddot{d}_s &= F_s - D_{ss}^{-1} \left[ W_s \dddot{r} + A_s^T (\dddot{\theta} + \omega_0 \dot{\theta}) + M_{ms}^T \dddot{d}_m + M_{ss} \dddot{d}_s + \\
+ D_{ms}^T \dddot{d}_m + K_{ms}^T \dddot{d}_m \right] \).
\)
\[ \ddot{d}_s = F_s - M_{ss}^{-1} [W_s^T \ddot{r} + A_s^T (\ddot{\theta} + \omega_0 \dot{\theta}) + M_{ms} \ddot{d}_m + D_{ms} \dot{d}_m + \\
D_{ss} \dot{d}_s + K_{ms}^T \ddot{d}_m + K_{ss} \ddot{d}_s ] , \] (3.26)

Substituting (3.24), (3.25) and (3.26) in (3.20), (3.21) and (3.22), results in:

\[ m \dddot{r} + Q (\dddot{\theta} + \omega_0 \dot{\theta}) + (W_m - W_s K_{ss}^{-1} K_{ms}^T) \ddot{d}_m = \\
= F_d + W_s \left[ M_{ss}^{-1} [W_s^T \dddot{r} + A_s^T (\dddot{\theta} + \omega_0 \dot{\theta}) + D_{ms} \dot{d}_m + \\
D_{ss} \dot{d}_s + K_{ms}^T \ddot{d}_m + K_{ss} \ddot{d}_s ] + (M_{ss}^{-1} M_{ms} - K_{ss}^{-1} K_{ms}^T) \ddot{d}_m - F_s \right] , \] (3.27)

\[ Q^T \dddot{r} + I (\dddot{\theta} + \omega_0 \dot{\theta}) + (A_m - A_s K_{ss}^{-1} K_{ms}) \ddot{d}_m = \\
= T_d + A_s \left[ M_{ss}^{-1} [W_s^T \dddot{r} + A_s^T (\dddot{\theta} + \omega_0 \dot{\theta}) + D_{ms} \dot{d}_m + \\
D_{ss} \dot{d}_s + K_{ms}^T \ddot{d}_m + K_{ss} \ddot{d}_s ] + (M_{ss}^{-1} M_{ms} - K_{ss}^{-1} K_{ms}^T) \ddot{d}_m - F_s \right] , \] (3.28)

\[ (W_m^T - K_{ms} K_{ss}^{-1} W_s) \dddot{r} + (A_m^T - K_{ms} K_{ss}^{-1} A_s^T) (\dddot{\theta} + \omega_0 \dot{\theta}) + \\
+ (M_{mm} - K_{ms} K_{ss}^{-1} M_{ms} - M_{ms} K_{ss}^{-1} K_{ms}^T + K_{ms} K_{ss}^{-1} M_{ss} K_{ms}^{-1} K_{ms}^T) \ddot{d}_m + \\
+ (D_{mm} - K_{ms} K_{ss}^{-1} D_{ms} - D_{ms} K_{ss}^{-1} K_{ms} + K_{ms} K_{ss}^{-1} D_{ss} K_{ms}^{-1} K_{ms}) \dot{d}_m + \\
+ (K_{mm} - K_{ms} K_{ss}^{-1} K_{ms}^T) \ddot{d}_m = \]
\[
\begin{align*}
= F_m + (M^{-1}_m + D^{-1}_m + D^{-1}_s) \left[ W^T \dot{r} + A^T (\dot{\phi} + \omega_0 \phi) + M^T \dot{d}_m \right. \\
+ D^T \dot{d}_m + K^T \dot{d}_m + K^{-1}_m d_m + K^{-1}_s d_s \left. \right] + (K^{-1}_m M^{-1}_s K^{-1}_s M^{-1}_m - M^{-1}_m K^{-1}_s K^T) d_m \\
+ (K^{-1}_m D^{-1}_s K^{-1}_s D^{-1}_m - D^{-1}_m K^{-1}_s K^T) d_m + (D^{-1}_m D^{-1}_s + K^{-1}_m K^{-1}_s) M^T \dot{d}_m \\
+ (M^{-1}_m + K^{-1}_m D^{-1}_s + D^{-1}_m + K^{-1}_s + K^{-1}_m) F_s .
\end{align*}
\] (3.29)

Comparing (3.27), (3.28) and (3.29) with (3.16), it is easily found that the left hand sides of (3.27), (3.28) and (3.29) coincide with (3.16), but extra terms exist on the right hand sides in (3.27), (3.28) and (3.29). It follows, that the reduced model described by (3.16) does not include the extra terms from an exact analysis. It means that in principle, static condensation leads to modelling errors.

Let us define a vector \( f \) in which all the extra terms in (3.27), (3.28) and (3.29) are included, An exact model of reduced order can then be obtained from (3.16), (3.27), (3.28) and (3.29) as:

\[
\dot{F}_r = A_r \dot{x}_r + B_r u_r + f ,
\] (3.30)

where \( f \) includes all the extra terms of (3.27), (3.28) and (3.29) as compared to (3.16), i.e.:
\[
\begin{align*}
W_s &= M_s^{-1} [W_s T s + A_s (\ddot{\theta} + \omega_0 \dot{\theta}) + D_{ms} \dot{d}_s + D_{ss} d_s + K_{ms} d_m + \\
&\quad + K_{ss} d_s] + (M_{ss}^{-1} M_{ms} - K_{ss} K_{ms}) d_m - F_s \\
A_s &= M_s^{-1} [W_s T s + A_s (\ddot{\theta} + \omega_0 \dot{\theta}) + D_{ms} \dot{d}_m + D_{ss} d_m + K_{ms} d_m + \\
&\quad + K_{ss} d_s] + (M_{ss}^{-1} M_{ms} - K_{ss} K_{ms}) d_m - F_s \\
f &= M_{ms}^{-1} + D_{ms}^{-1} [W_s T s + A_s (\ddot{\theta} + \omega_0 \dot{\theta}) + M_{ms} \dot{d}_m + D_{ms} d_m + \\
&\quad + K_{ms} d_m + K_{ss} d_s] + (K_{ms}^{-1} M_{ss} K_{ms}^{-1} K_{ms} T s - M_{ms} K_{ss} K_{ms}) d_m + \\
&\quad + (K_{ms}^{-1} D_{ss} K_{ms}^{-1} K_{ms} T s - D_{ms} K_{ss} K_{ms}) d_m - (M_{ms} + D_{ms} + K_{ms}) F_s + \\
&\quad + (D_{ms}^{-1} + K_{ms} K_{ss}^{-1}) M_{ss} \ddot{d}_s + (M_{ms} M_{ss}^{-1} + K_{ms} K_{ss}^{-1}) D_{ss} \dot{d}_s \\
&\quad + \omega_0 m_r
\end{align*}
\]

(3.31)

In (3.30) the original high order dynamical model of a flexible spacecraft, described by Eq. (2.91), is replaced by a reduced order model with an additional input vector \( f \) in (3.31). The latter describes the modelling errors resulting from order reduction. If only reduced order models are used, the modelling error vector \( f \) is in principle unknown. If the reduced order model (3.30) is used for parameter estimation, the modelling error vector \( f \) should be estimated as part of the model. In this study, the modelling error vector \( f \) in (3.31) is replaced by some other mathematical representation, in order to avoid direct calculation of the vector \( f \) using (3.31). This mathematical representation may be a static or a dynamic model. The unknown parameters of this latter model will be estimated simultaneously with the parameters in
the reduced order model (see next section). A more detailed discussion of the modelling error vector \( \bar{f} \) is given in the next subsection.

As mentioned above, the measured displacements are selected as the master displacements to be retained in the reduced order system model, so the observation equation for the reduced system model (3.30) can now be written as:

\[
Y_r = C_{11} \dot{x}_r + C_{21} x_r + \bar{v}_1,
\]

(3.32)

where \( C_{11} \) is the observation matrix related to the measurements of the specific forces of the reduced order system and \( C_{21} \) is the observation matrix related to the measurements of the velocities and displacements of the rigid main body of the spacecraft.

Referring to the reduced order state vector \( \bar{x}_r \) described in (3.16), the matrix \( C_{11} \) should have the form:

\[
C_{11} = \begin{bmatrix}
C_{q_a x(6+m_r)} & 0_{q_a x(6+m_r)} \\
0_{(q-q_a)x(6+m_r)} & 0_{(q-q_a)x(6+m_r)}
\end{bmatrix},
\]

(3.33)

where \( C_{q_a x(6+m_r)} \) is a \( q_a \times (6 + m_r) \) nonzero matrix related to the measurements of specific force variables of the spacecraft.

The matrix \( C_{21} \) can be written as:

\[
C_{21} = \begin{bmatrix}
0_{q_a x3} & 0_{q_a x3} & 0_{q_a x3} & 0_{q_a x3} & 0_{q_a x3} & 0_{q_a x3} \\
C_{q_v x3} & 0_{q_v x3} & 0_{q_v x3} & 0_{q_v x3} & 0_{q_v x3} & 0_{q_v x3} \\
0_{q_w x3} & C_{q_w x3} & 0_{q_w x3} & 0_{q_w x3} & 0_{q_w x3} & 0_{q_w x3} \\
0_{q_r x3} & 0_{q_r x3} & C_{q_r x3} & 0_{q_r x3} & 0_{q_r x3} & 0_{q_r x3} \\
0_{q_\theta x3} & 0_{q_\theta x3} & 0_{q_\theta x3} & C_{q_\theta x3} & 0_{q_\theta x3} & 0_{q_\theta x3} \\
\end{bmatrix}.
\]

(3.34)
In (3.34) \( C_{q_v} \) is the nonzero observation submatrix related to the velocity measurements of the rigid main body, \( C_{q_w} \) is the nonzero observation submatrix related to the angular velocity measurements of the rigid main body, \( C_{q_r} \) is the nonzero observation submatrix related to the translational displacement measurements of the rigid main body and \( C_{q_\theta} \) is the nonzero observation submatrix related to the rotational displacement measurements of the rigid main body. The first row of (3.34) is zero, since in (3.33) the specific force measurements of the flexible substructures as well as the rigid main body have been included. In practice, no measurements are made of rigid body translational velocities and displacements, i.e.:

\[
q_v = 0 \quad , \quad (3.35)
\]

and:

\[
q_r = 0 \quad . \quad (3.36)
\]

The observation matrix \( C_z \) may now be simplified as follows:

\[
C_z = \begin{bmatrix}
0 q_x 3 & 0 q_x 3 & 0 q_x m_r & 0 q_x 3 & 0 q_x 3 & 0 q_x m_r \\
0 q_\omega 3 & 0 q_\omega 3 & 0 q_\omega m_r & 0 q_\omega 3 & 0 q_\omega 3 & 0 q_\omega m_r \\
0 q_\theta 3 & 0 q_\theta 3 & 0 q_\theta m_r & 0 q_\theta 3 & 0 q_\theta 3 & 0 q_\theta m_r \\
\end{bmatrix} . \quad (3.37)
\]

Since (3.30) can also be written as:

\[
\dot{x}_r = F^{-1}_r A_r x_r + F^{-1}_r B_r u + F^{-1}_r f' \quad , \quad (3.38)
\]

an alternate form of the output equation of the reduced order system model can be written, by substituting (3.38) in (3.32), as:
\[ \mathbf{y}_r = C_{=1_1} (E_{=1_1}^{-1} \mathbf{A}_r \mathbf{x}_r + F_{=1_1}^{-1} \mathbf{B}_r \mathbf{u} + F_{=1_1}^{-1} \mathbf{f}) + C_{=2_1} \mathbf{x}_r + \mathbf{y}_1 \]

\[ = H_{=1_1} \mathbf{x}_r + D_{=1_1} \mathbf{u} + \mathbf{y}_1, \tag{3.39} \]

where:

\[ H_{=1_1} = C_{=1_1} F_{=1_1}^{-1} \mathbf{A}_r + C_{=2_1}, \tag{3.40} \]

\[ D_{=1_1} = C_{=1_1} F_{=1_1}^{-1} \mathbf{B}_r, \tag{3.41} \]

\[ \mathbf{y} = C_{=1_1} F_{=1_1}^{-1} \mathbf{f} + \mathbf{y}_1. \tag{3.42} \]

Summarizing, the reduced system model outputs \( \mathbf{y}_r \) must be expected to be different from the original system model outputs \( \mathbf{y} \), described in (3.1). That is, \( \mathbf{y}_r \neq \mathbf{y} \). In principle, if the modelling error vector \( \mathbf{f} \) is included as an additional input in the reduced order system model, and if it is assumed to be exactly equal to the extra terms described by (3.31), then \( \mathbf{y}_r = \mathbf{y} \). In the present study, \( \mathbf{f} \) is replaced by certain mathematical approximations, so that hopefully, this results in sufficiently small discrepancies between \( \mathbf{y}_r \) and \( \mathbf{y} \), according to some suitable criterion. This will be discussed in the next section.

### 3.2.3 Modelling error characteristics of the reduced model

In Subsection 3.2.2, the modelling errors of the reduced order system model were described by an input vector \( \mathbf{f} \). In this subsection this vector will be discussed in some detail.

In the field of applied control technology, it is a well known fact that all realistic control systems operate in environments that produce system disturbances of one kind or another. Here, the term 'disturbances' refers to
that special category of system inputs which are not accurately known before
hand and which can not be manipulated by the control designer, e.g., the
unknown inputs \( f \) in our case belong to that category.

Disturbances are an important factor in control design problems, because
they usually introduce unwanted disruptions in the otherwise orderly be-
haviour of the controlled system.

On the other hand, parameter estimation, state reconstruction as well as
system control design require these disturbances to be taken into account to
prevent the various algorithms from diverging [18].

The only unknown disturbances \( f \) considered in this study so far, are the
modelling errors resulting from model order reduction. In fact, several ex-
ternal disturbances must be expected to act on the spacecraft in orbit, for
example, aerodynamic disturbance forces and torques, solar pressure forces
and torques, disturbance forces and torques of the magnetic field of the
earth or other planets, disturbance forces and torques of the gravity field
of the earth, etc. These external disturbances may be taken into account by
assuming that the disturbance vector \( f \) contains both the what can be called
"internal disturbances" resulting from the system model order reduction, and
the external disturbances.

The kinds of disturbances \( f \) which one encounters in realistic systems can be
classified into two broad categories, see [9]:

1) noise-type disturbances and

2) disturbance with waveform structures.

Time recordings of noise-type disturbances are essentially ragged and er-
ragic in nature, having no significant degree of smoothness or regularity in
their waveforms. On the other hand, time recordings of disturbances which
have a 'waveform structure' exhibit distinguishable waveform patterns, at
least over the short time intervals of the recording.

Noise-type disturbances, which by the preceding definition have no sig-
nificant degree of waveform structure, are best characterized in terms of
their statistical properties such as mean value, covariance, power spectral
density, etc. In this way, one can mathematically model noise-type distur-
bance by traditional random processes, utilizing the notions of 'white
noise', 'colored noise', etc. The fields of stochastic stability, control,
and filtering theory are concerned almost entirely with noise-type disturbances of this latter variety. A number of excellent texts on those subjects are available, e.g., [6], [7] and [8].

In present study it is assumed, that the unknown inputs \( f \) acting on the reduced order system model can be interpreted as noise-type disturbances and adequately be represented in terms of random processes.

The reduced order model represents the low frequency region of the actual system response, as discussed in the the last subsection, so the disturbances caused by order reduction should represent the behaviour of the actual system in high frequency region. According to [9] and [39], such disturbances, in general, can be modelled in terms of a linear system model as:

\[
\dot{z} = \varepsilon_z z + \varepsilon_w w_2, \tag{3.43}
\]

\[
f = G_r z + G_1 w_1, \tag{3.44}
\]

where \( w_1 \) and \( w_2 \) denote independent white noise process vectors with zero mean and Gaussian distribution, \( z \) is a state variable vector which describes the filtered white noises, \( \varepsilon \) represents a state system matrix and \( G_r \) and \( G_1 \) represent the constant output matrices. If the disturbance vector \( f \) only contains the modelling errors resulting from order reduction, see (3.31) and (3.38), \( G_r \) and \( G_1 \) have the following form:

\[
G_{r} = \begin{bmatrix}
  G_r^* \\
  \frac{0}{(6+m_r)z} \\
\end{bmatrix}, \tag{3.45}
\]

\[
G_1 = \begin{bmatrix}
  G_1^* \\
  \frac{0}{(6+m_r)z} \\
\end{bmatrix}, \tag{3.46}
\]
where $G_r^*$ and $G_1^*$ are $(6 + m_r) \times m_z$ nonzero matrices.

Eqs. (3.43) and (3.44) show that the actual disturbance vector $f$ is replaced by a linear combination of the response of a dynamic system excited by white noises and purely white noise. Therefore, the reduced order system has turned into a stochastic system.

In the literature, Eq. (3.43) is also called a shaping filter. When (3.44) is used to represent the input disturbances, the reduced order system model as described by (3.30) and (3.35) is extended by the dynamical shaping filter (3.43). The augmented model with noise inputs has the following form:

$$
F_a^* x = A_a x_a + B_a u + G_a w, \quad (3.47)
$$

where:

$$
F_a = \begin{bmatrix}
F_r & 0 & 0 \\
0 & (12+2m_r)xm_z & 0 \\
0 & 0 & U = m_z
\end{bmatrix}, \quad (3.48)
$$

$$
A_a = \begin{bmatrix}
A_r & G_r \\
0 & (12+2m_r)E \\
0 & 0 & U = m_z
\end{bmatrix}, \quad (3.49)
$$

$$
G_a = \begin{bmatrix}
G_1 & 0 & 0 \\
0 & (12+2m_r)xm_z & 0 \\
0 & 0 & U = m_z
\end{bmatrix}, \quad (3.50)
$$

$$
B_a = \begin{bmatrix}
B_r \\
0 & 0 & 0
\end{bmatrix}, \quad (3.51)
$$

$$
x_a = [x_r^T | z^T]^T, \quad (3.52)
$$
\[ w = [w_1^T \mid w_2^T]^T. \quad (3.53) \]

As mentioned above, the input vectors \( w_1 \) and \( w_2 \) are white, zero-mean random processes with independent and Gaussian distributions. The statistical properties of \( w_1 \) and \( w_2 \) are defined as:

\[ E\{w(t)\} = 0, \quad (3.54) \]

\[ E\{w(t) w^T(\tau)\} = V_{ww} \delta(t - \tau), \quad (3.55) \]

where:

\[ V_{ww} = \begin{bmatrix} V_{w1} & 0_{m_1 \times m_2} \\ 0_{m_2 \times m_1} & V_{w2} \end{bmatrix}, \quad (3.56) \]

in which \( V_{ww} \) denotes the covariance matrix of the white noises \( w_1 \) and \( w_2 \), \( \delta(t - \tau) \) is the Dirac delta function and \( V_{w1} \) and \( V_{w2} \) are the not necessarily diagonal covariance submatrices of \( w_1 \) and \( w_2 \), respectively.

Eq. (3.47) shows, that the system model with unknown disturbances caused by order reduction, is now represented by a stochastic dynamical system model. The system model contains additional parameters describing the statistic properties of the disturbing random noise, see (3.56).

Referring to (3.39) and (3.47), the system model output equation can be written as:

\[
y = (C_1 \mid 0_{q \times m}) x_a + (C_2 \mid 0_{q \times m}) x_a + y_i \\
= H x_a + D u + v, \quad (3.54)
\]
where:

\[
H = \begin{bmatrix}
    C_1 F_r^{-1} A_r + C_2 & C_1 F_r^{-1} G_r
\end{bmatrix},
\]

(3.55)

\[
D = C_1 F_r^{-1} B_r,
\]

(3.56)

\[
v = (C_1 F_r^{-1} G_1 \mid O_{q \times m_z}) w + v_1,
\]

(3.57)

\(H\) and \(D\) are called the observation matrix and feed-forward matrix, respectively.

Here is assumed, that the measurement noise \(v_1\) is independent of the process noises \(w\), so that the statistical properties of \(v_1\), \(v\) and \(w\) can be obtained as follows:

\[
E\{v_1(t)\} = 0_q,
\]

(3.58)

\[
E\{v_1(t) v_1^T(\tau)\} = V_{v_1} \delta(\tau - t),
\]

(3.59)

\[
E\{w(t) v_1(\tau)\} = O_{(m_w + m_z) \times q},
\]

(3.60)

\[
E\{v_1(t)\} = (C_1 F_r^{-1} G_1 \mid O_{q \times m_z}) E\{w(t)\} + E\{v_1(t)\} = 0_q,
\]

(3.61)

\[
E\{v(t) v_1^T(\tau)\} = E\{[(C_1 F_r^{-1} G_1 \mid O_{q \times m_z}) w(t) + v_1(t)]
\]

\[
\times [(C_1 F_r^{-1} G_1 \mid O_{q \times m_z}) w(\tau) + v_1(\tau)]^T]\]

\[
= V_{v} \delta(\tau - t),
\]

(3.62)
In (3.62) it follows that:

\[ W = C_1 F^{-1} G_1 V W \ T = C_1 F^{-T} C_1^T + V_1. \] (3.63)

The covariance matrix of \( w(t) \) and \( y(\tau) \) can be derived as:

\[ E \{ w(t) y^T(\tau) \} = E \{ w(t) \left[ \begin{array}{ccc} C_1 & G_1 & 0 \\ F^{-1} & G_1 & 0 \\ 0 & 0 & o \end{array} \right] y(\tau) + V_1(\tau) \} \]

\[ = V_{wV} \delta(\tau - t), \] (3.64)

where:

\[ V_{wV} = V_{w1} C_1 F^{-T} C_1. \] (3.65)

In this subsection, the disturbance inputs \( f \) have been discussed in some detail. The disturbance input vector \( f \) has been represented by a combination of a dynamical linear system with white noise inputs (shaping filter), and white noises, see Eqs. (3.43) and (3.44). The purpose of the present subsection is to present an adequate model of the modelling errors as, caused by the system model order reduction. In the next section, the so-called maximum likelihood estimator will be introduced which is used to identify the parameters in the resulting stochastic system model.

### 3.2.4 Discretization of the system model

Eqs. (3.47) and (3.54) represent a model of a flexible spacecraft in continuous time. The measurements of system inputs and outputs, however, are only available in discrete time. It is appropriate, therefore, to develop the discrete-time form of (3.47) and (3.54).

A continuous-time dynamical system represented by a set of differential equations, see Eq. (3.47), can be approximated by a set of difference equations (see, e.g., [25]).
A linear discrete-time system in state space can be written in the form of vector-matrix difference equations as follows:

\[
\mathbf{x}_a[(k+1)T] = \Phi(T) \mathbf{x}_a(kT) + \Gamma_u(T) \mathbf{u}(kT) + \Gamma_w(T) \mathbf{w}(kT),
\]

where \( T \) is the discrete time interval, \( \Phi(T) \) is the transition matrix defined by:

\[
\Phi(T) = \exp\left[\mathbf{F}_a^{-1} \mathbf{A}_a T\right],
\]

where \( \mathbf{F}_a \) and \( \mathbf{A}_a \) denote the system matrices of the original continuous-time system (3.47).

\( \Gamma_u(T) \) and \( \Gamma_w(T) \) in (3.66) are the deterministic input distribution matrix and the noise input distribution matrix respectively, i.e.:

\[
\Gamma_u(T) = \int_0^T \Phi(\tau) \, d\tau \mathbf{F}_a^{-1} \mathbf{B}_a,
\]

\[
\Gamma_w(T) = \int_0^T \Phi(\tau) \, d\tau \mathbf{F}_a^{-1} \mathbf{G}_a.
\]

in which \( \mathbf{B}_a \) and \( \mathbf{G}_a \) denote the deterministic and stochastic input matrices of the original continuous-time system (3.47).

Here it is assumed that during the time interval \( T \), \( \mathbf{u}(kT) \) and \( \mathbf{w}(kT) \) are constant.

There are many methods to calculate the transition matrix \( \Phi(T) \) (see Ref. [25]). One way is to use the power series expansions in \( \mathbf{F}_a^{-1} \mathbf{A}_a T \), i.e.:

\[
\Phi(T) = \exp\left[\mathbf{F}_a^{-1} \mathbf{A}_a T\right] = I + \sum_{n=1}^{\infty} \frac{1}{n!} \left[\mathbf{F}_a^{-1} \mathbf{A}_a\right]^n T^n.
\]
Substituting (3.70) in (3.68) and (3.69), the matrices \( \Gamma_u(T) \) and \( \Gamma_w(T) \) are obtained as follows:

\[
\Gamma_u(T) = T \left[ U + \sum_{n=1}^{\infty} \frac{1}{(n+1)!} \left( F_a^{-1} A_a \right)^n T^n \right] \left( F_a^{-1} B_a \right), \tag{3.71}
\]

\[
\Gamma_w(T) = T \left[ W + \sum_{n=1}^{\infty} \frac{1}{(n+1)!} \left( F_a^{-1} A_a \right)^n T^n \right] \left( F_a^{-1} G_a \right). \tag{3.72}
\]

In the following analysis, the sampling time \( T \) as well as the subscript \( a \) in the discrete-time system will be dropped for simplicity, i.e.:

\[
x(k+1) = \Phi x(k) + \Gamma_u u(k) + \Gamma_w w(k), \tag{3.73}
\]

\[
y(k) = H x(k) + D u(k) + v(k). \tag{3.74}
\]

Referring to Subsection 3.2.3, the statistical properties of \( w(k) \) and \( v(k) \) are given as follows:

\[
E\{w(k)\} = 0_{m_w + m_z}, \tag{3.75}
\]

\[
E\{v(k)\} = 0_q, \tag{3.76}
\]

\[
E\{w(k) w^T(i)\} = V_w \delta_{k,i}, \tag{3.77}
\]

\[
E\{v(k) v^T(i)\} = \left( C_1 F_r^{-1} G_1 V_{w1} G_1^T F_r^{-T} C_1^T + V_{v1} \right) \delta_{k,i}, \tag{3.78}
\]

\[
E\{w(k) v^T(i)\} = V_{w1} G_1^T F_r^{-T} C_1 \delta_{k,i}. \tag{3.79}
\]

where \( \delta_{k,i} \) denotes the Kronecker delta function.
3.2.5 Parameters to be identified in the total system

All the preceding discussions focus on the mathematical model of flexible spacecraft to be used for parameter identification. The system model (3.73) through (3.79) contains the parameters to be identified. In Chapter 2, some structural parameters have been defined, e.g., the dimensions of the finite elements (a, b, t), the Young's modulus (E), the Poisson ratio (ν), the mass densities of the elements (ρ), the damping coefficients (α, β), the total mass of the spacecraft (m), the total inertia matrix (I), the coupling matrix Q, etc.

When the original high order system model described in Chapter 2 is reduced to the model represented by (3.73) through (3.79), the structural parameters mentioned above remain unchanged (see Subsection 3.2.3). There are some more parameters which are contained in the input matrix B, the observation matrix H, the feedforward matrix D, the shaping filter matrix E and the covariance matrices \( V_{wv} \), \( V_v \) and \( V_w \). These parameters should be identified as well. Now a total parameter vector \( \theta \) may be defined as:

\[
\theta = [\theta_P^T \theta_B^T \theta_H^T \theta_D^T \theta_E^T \theta_w^T \theta_v^T \theta_{wv}^T]^T
\]  

(3.80)

The parameter subvectors in \( \theta \) are defined as follows:

- \( \theta_P \) - contains all the structural parameters of Chapter 2, i.e., \( \theta_P = P \).
- \( \theta_B \) - contains all the unknown parameters in the deterministic input matrix B (e.g., the parameters of the actuators on the spacecraft).
- \( \theta_H \) - contains the parameters in the observation matrix H not included in \( \theta_P \) and \( \theta_B \), i.e. the unknown parameters in \( C_1 \) and \( C_2 \), see(3.55).
- \( \theta_D \) - contains all the parameters in the feedforward matrix D not included in \( \theta_P \), \( \theta_B \) and \( \theta_H \). This means that in fact, \( \theta_D \) is empty, see (3.56).
- \( \theta_E \) - contains all the parameters in the shaping filter matrix E. This matrix will be discussed in the next chapter.
\( \theta_w \) contains all the parameters in the covariance matrix of the process noises \( \Sigma_w \).

\( \theta_v \) contains all the unknown parameters in the covariance matrix \( \Sigma_v \) of the measurement noises not yet mentioned above. It follows from (3.78), that it contains only the unknown parameters in covariance matrix \( \Sigma_v' \).

\( \theta_{wv} \) contains all the parameters in the covariance matrix of the correlated process and measurement noises, i.e. \( \Sigma_{wv} \), not yet mentioned above. It follows from (3.79) that \( \theta_{wv} \) is empty.

### 3.3 Maximum Likelihood Parameter Estimation of Linear Systems with Correlated Process and Measurement Noises

In this section, parameter estimation schemes for estimating the parameters in a stochastic system, as described in the last section, are discussed.

In parameter estimation, some objective criterion is commonly used to measure the relative merits of one set of estimated parameters compared to another. For example, the criterion could be the sum of squares of the difference between the model outputs and actual system outputs over some period of time. In some cases, the criterion may be directly related to the intended application. For example, if one is interested in the d-step-ahead prediction error, the criterion could be the mean-square d-step-ahead prediction error [26]. In other cases, the choice of the criterion may be also motivated by statistical considerations in an effort to arrive at a 'good' description of the system. Examples of this latter type of criterion are the maximum likelihood and Bayesian criteria (Goodwin and Payne, 1977, Astrom and Eykhoff, 1971). In this study, the maximum likelihood criterion is used. In particular, the application of maximum likelihood parameter estimation algorithms, to determine the parameters in our stochastic system with correlated process and measurement noises (see Eq. (3.73) through (3.79)), is considered. The reason to choose the maximum likelihood criteria has been given in Chapter 1, so that here only the algorithms for our particular model are discussed.
The basic idea of the maximum likelihood method is to construct a function of the data and the unknown parameters, which is called the likelihood function. If a sequence of observations \( \vec{X}(1), \ldots, \vec{X}(N) \) is made of the disturbed system states (see Eq. (3.73) and (3.74)), the maximum likelihood estimate of the unknown parameter vector \( \theta \) (3.80) is obtained by maximizing the conditional probability density function \( p[\vec{Y}_N', \theta] \), i.e.:

\[
\hat{\theta} = \max_\theta p[\vec{Y}_N', \theta],
\]

where:

\[
\vec{Y}_N = [\vec{Y}(1|0), \vec{Y}(2|1), \ldots, \vec{Y}(k+1|k), \ldots, \vec{Y}(N|N-1)].
\]

In (3.82), \( \vec{Y}(k+1|k) \) is called the prediction error vector (or innovation vector) (see e.g. [12], [16] and [19]), It is defined by:

\[
\vec{Y}(k+1|k) = \vec{X}(k+1) - \hat{\vec{X}}(k+1|k),
\]

where \( \vec{X}(k+1) \) is the observation vector (3.74) at sample time \( k+1 \) and \( \hat{\vec{X}}(k+1|k) \) is the predicted output vector as generated by a Kalman filter (see e.g. [16], [18], [28] and [29]).

If \( \vec{X}(k+1), k = 0, 1, 2, \ldots, N-1 \), are all independent, an expression for \( p[\vec{Y}_N', \theta] \) can be derived from a successive application of Bayes rule as:

\[
p[\vec{Y}_N', \theta] = p[\vec{Y}(1|0), \vec{Y}(2|1), \ldots, \vec{Y}(k+1|k), \ldots, \vec{Y}(N|N-1), \theta]
\]

\[
= p[\vec{Y}(N|N-1), \vec{Y}_{N-1}, \theta] p[\vec{Y}_{N-1}, \theta]
\]
\[
= p[\tilde{y}(N|N-1), \tilde{y}_{N-1}, \theta] p[\tilde{y}(N-1|N-2), \tilde{y}_{N-2}, \theta] p[\tilde{y}_{N-2}, \theta]
\]

\[
\prod_{k=0}^{N-1} p[\tilde{y}(k+1|k), \tilde{y}_{k-1}, \theta] .
\]

The Gaussian probability density function \( p[\tilde{y}(k+1|k), \tilde{y}_{k-1}, \theta] \) is:

\[
p[\tilde{y}(k+1|k), \tilde{y}_{k-1}, \theta] = (2\pi)^{-q/2} \det [\Sigma_e(k+1|k)]^{-1/2} \times
\]

\[
x \exp \left[-\frac{1}{2} \tilde{y}^T(k+1|k) \Sigma_e^{-1}(k+1|k) \tilde{y}(k+1|k) \right] ,
\]

\[(3.85)\]

where \( \Sigma_e(k+1|k) \) is called the covariance matrix of prediction errors (or innovations), i.e.:

\[
\Sigma_e(k+1|k) = \mathbb{E}[\tilde{y}(k+1|k) \tilde{y}^T(k+1|k)] .
\]

\[(3.86)\]

The conditional probability density function \( p[\tilde{y}_N, \theta] \) in \((3.81)\) is called the likelihood function. Substituting \((3.85)\) in \((3.84)\), it results in:

\[
p[\tilde{y}_N, \theta] = \prod_{k=0}^{N-1} \left[(2\pi)^{-q/2} \det [\Sigma_e(k+1|k)]^{-1/2} \times
\]

\[
x \exp \left[-\frac{1}{2} \tilde{y}^T(k+1|k) \Sigma_e^{-1}(k+1|k) \tilde{y}(k+1|k) \right] .
\]

\[(3.87)\]

It is convenient to work with the logarithm of the likelihood function \((3.87)\) for computational purposes. The negative of the logarithm of the likelihood function is used, to conform to optimization conventions which
work with minimization rather than maximization (see e.g. [15], [16] and
[19]). Ignoring the unnecessary constant term, the negative logarithm of
the likelihood function is:

\[
L(\theta) = \frac{1}{2} \sum_{k=0}^{N-1} \ln \det [\Sigma_e(k+1|k)] + \\
+ \frac{1}{2} \sum_{k=0}^{N-1} [\tilde{\Sigma}(k+1|k) \Sigma_e^{-1}(k+1|k) \tilde{\Sigma}(k+1|k)].
\] (3.88)

The covariance matrix of the prediction errors \(\Sigma_e(k+1|k)\) can be obtained as
follows.
Since the prediction errors can be written as:

\[
\tilde{\Sigma}(k+1|k) = y(k+1) - \hat{x}(k+1|k)
\]

\[
= H [\tilde{x}(k+1) - \hat{x}(k+1|k)] + y(k+1),
\] (3.89)

where \(\hat{x}(k+1|k)\) is the one stage prediction of states, the products of
\(\tilde{\Sigma}(k+1|k)\) and \(\tilde{\Sigma}(k+1|k)^T\) can be written as:

\[
\tilde{\Sigma}(k+1,k) \tilde{\Sigma}(k+1,k)^T = H [\tilde{x}(k+1) - \hat{x}(k+1|k)] [\tilde{x}(k+1) - \hat{x}(k+1|k)]^T H^T + \\
+ H [\tilde{x}(k+1) - \hat{x}(k+1|k)] \Sigma_e^{-1}(k+1) + \\
+ y(k+1) [\tilde{x}(k+1) - \hat{x}(k+1|k)]^T H^T + y(k+1) \Sigma_e^{-1}(k+1) \Sigma_e^{-1}(k+1)^T.
\] (3.90)

Taking the expectation of \(\tilde{\Sigma}(k+1|k) \tilde{\Sigma}(k+1|k)^T\), the result is:
\[ V_{e}(k+1|k) = H P(k+1|k) H^{T} + V_{v}, \]  

(3.91)

where \( P(k+1,k) \) is the covariance matrix of the predicted errors of the states, see Eq. (3.97).

In (3.91) the following definitions from Subsection 3.2.4 were used:

\[ E \{ y(k) \} = 0, \]

\[ E \{ y(k) y^{T}(k) \} = V_{y}. \]

The Kalman filter provides a way of estimating the states \( \hat{x}(k+1) \) of the model (3.73) to (3.79). The filter has the following two interpretations (see [18], [26], [39] and [40]).

1) If the noises are gaussian, the filter gives the minimum variance estimates of the states, i.e. the conditional mean of \( \hat{x}(k+1) \) given the past data \( \{ y(k), y(k-1), ... \} \).

2) If the gaussian assumption is removed, the filter gives the linear minimum variance estimates of the states (i.e., having the smallest unconditional error covariance among all linear estimates, but this will not, in general, be the conditional mean).

For simplicity, the filter will be derived under the assumption that the noises are gaussian. It is pointed out that the algorithms described here are not unique, as there is an alternate formulation of the Kalman filter (see e.g. [39] and [40]). The maximum likelihood parameter estimator with this latter formulation is given in Ref. 3.45. Then the following results are obtained.

Considering the system of (3.73) to (3.79) and assuming that the initial state and noise sequences are jointly gaussian, the Kalman filter of the linear system with correlated process and measurement noises (see Eq. (3.79)) can be found from [18] and [40] as follows.

The one stage prediction algorithm is given by:
\[ \hat{x}(k+1|k) = \Phi \hat{x}(k|k) + \Gamma_u u(k) + K_p [y(k) - \hat{y}(k|k)] \] 

where:

\[ \hat{y}(k|k) = \hat{H} \hat{x}(k|k) + D u(k) \] 

and \([y(k) - \hat{y}(k|k)]\) is called the residue vector of the outputs, and \(K_p\) is called the one stage prediction gain:

\[ K_p = \Gamma_w W_{\hat{V}} V_{\hat{V}}^{-1} \] 

The measurement update of the filter is:

\[ \hat{x}(k+1|k+1) = \hat{x}(k+1|k) + K(k+1) \hat{y}(k+1|k) \] 

where \(K(k+1)\) is the Kalman filter gain:

\[ K(k+1) = P(k+1|k) H^T \left[ H P(k+1|k) H^T + V_{\hat{V}} \right]^{-1} \] 

The covariance matrix of the predicted errors of the state as mentioned in (3.91) has the form:

\[ P(k+1|k) = E\left[ [\hat{x}(k+1) - \hat{x}(k+1|k)] [\hat{x}(k+1) - \hat{x}(k+1|k)]^T \right] \] 

and it can be calculated as:

\[ P(k+1|k) = (\Phi - K_p \hat{H}) P(k|k) (\Phi - K_p \hat{H})^T + \Gamma_w W_{\hat{V}} V_{\hat{V}} - K_p V_{\hat{V}} K_p^T \] 

in which \(P(k|k)\) is the a posteriori covariance matrix of the state estimates, i.e.:
\[\hat{P}(k+1|k+1) = E\left[(\hat{x}(k+1) - \hat{x}(k+1|k+1)) [\hat{x}(k+1) - \hat{x}(k+1|k+1)]^T\right]. \quad (3.99)\]

From [18], (3.99) can be obtained as:

\[\hat{P}(k+1|k+1) = \hat{P}(k+1|k) - \hat{K}(k+1) H \hat{P}(k+1|k). \quad (3.100)\]

With (3.91), the Kalman gain described by (3.96) can also be written as:

\[\hat{K}(k+1) = \hat{P}(k+1,k) H^T \hat{V}^{-1}(k+1,k). \quad (3.96b)\]

The problem of determining the maximum likelihood estimates of the unknown parameters \(\hat{\theta}\) of (3.81) has now become one of finding a way to minimize the negative logarithm of the likelihood function described by (3.88).

The necessary condition for the optimum of the likelihood function is given by (see e.g. [10], [11], [12], [13], [14], [15], [16], [17], [18] and [19]):

\[\frac{\partial}{\partial \hat{\theta}} [L(\hat{\theta})] \left|_{\hat{\theta} = \hat{\theta}_{ML}} = 0. \quad (3.101)\]

where \(\hat{\theta}_{ML}\) is the maximum likelihood estimate of \(\hat{\theta}\).

Expression (3.101) is called the likelihood equation. The results of this set of equations yielding the smallest value of \(L(\hat{\theta})\), are the maximum likelihood estimates of the unknown parameters \(\hat{\theta}_{ML}\).

### 3.4 Optimization Methods for ML Parameter Estimation

In principle, the negative logarithm of the likelihood function given by (3.88) may be minimized with respect to the parameter vector \(\hat{\theta}\) by a direct search method, which uses the likelihood function values only, e.g., Powell's method, the conjugate direction method, the random search method and the method of Hooke and Jeeves (see e.g. [30], [31], [32], [33], [34], [35] and
These methods have in common, that no gradient of the criterion function (the negative logarithm of the likelihood function in our case) needs to be calculated. In principle this is an important advantage, since the calculation of the gradient of the criterion function, i.e., \( \frac{\partial}{\partial \theta} [L(\theta)] \) (3.101) may be very cumbersome, due to the complexity of the likelihood function.

However, convergence rates of these optimization methods may be poor (see e.g. [30], [32], [35] and [36]). Usually, more efficient algorithms result if the gradient of the criterion function with respect to the parameters is taken into account. This approach is taken in so-called gradient methods (see e.g. [30], [32], [35] and [36]), where the parameter vector \( \hat{\theta} \) is adjusted in each iteration \( i \) by:

\[
\hat{\theta}(i+1) = \hat{\theta}(i) + \alpha(i) \mathbf{g}(i),
\]

where \( \hat{\theta}(i+1) \) is the parameter vector \( \hat{\theta} \) in the \((i+1)\)th iteration, starting from the parameter vector in the \( i \)th iteration \( \hat{\theta}(i) \), \( \alpha(i) \) is a factor determined by a one-dimensional search for the minimum of the negative logarithm of the likelihood function \( L(\hat{\theta}) \) and \( \mathbf{g}(i) \) is the gradient vector of the negative logarithm of the likelihood function, i.e.:

\[
\mathbf{g}(i) = \left. \frac{\partial}{\partial \theta} [L(\theta)] \right|_{\theta = \hat{\theta}(i)}.
\]

Gradient methods only make use of the first order gradient of the criterion function, e.g. the method of steepest descent and the method of conjugate gradients (see e.g. [30], [32], [35] and [36]).

According to the literature, a much better convergence results if the Newton-Raphson method is applied. In the Newton-Raphson method the parameter vector \( \hat{\theta} \) is adjusted according to:
\[ \theta(i+1) = \theta(i) + \alpha(1) R^{-1}(1) \tilde{g}(1) \]  

(3.104)

in which \( R(i) \) is a matrix of second order partial derivatives called the Hesse matrix, see [15], [16] and [19], i.e.:

\[ R(i) = \frac{\partial^2}{\partial \theta^2}[L(\theta)] \bigg| \theta = \theta(1) \]  

(3.105)

In this method, the second order gradient of the negative logarithm of the likelihood function with respect to the parameter vector \( \theta \) is taken into account. If the Newton-Raphson method is convergent, the convergence is quadratic. However, the method has the following drawbacks, see [15], [16] and [19]:

1) It fails to converge to the desired optimum, whenever the Hesse matrix has some negative eigenvalues.

2) If the Hesse matrix is nearly singular, there are numerical problems in inverting it. This may result in a slow convergence, or in no convergence at all.

3) Generally, the computation of the exact Hesse matrix is rather time consuming, since second order gradients of the state estimator described by (3.92) through (3.100) must be evaluated.

For the above reasons, the Newton-Raphson method is generally not used in parameter estimation problems, but rather the so-called Gauss-Newton method in which the Hesse matrix \( R(i) \), as defined by Eq. (3.105), is replaced by an approximation matrix, i.e. the so-called Fisher information matrix, composed of only the first order gradients of the criterion function.

The Fisher information matrix (see e.g. [15], [16], [19] and [26]) is defined as:

\[ R(i) = E \left\{ \frac{\partial^2}{\partial \theta^2}[L(\theta)] \right\} \bigg| \theta = \theta(1) \]  

(3.106a)

or, alternatively written as:
\[ R(i) = \mathbb{E} \left\{ \frac{\partial}{\partial \theta} [L(\theta)] \mid \frac{\partial}{\partial \theta} ^T [L(\theta)] \right\} \mid \theta = \theta(i) \]  

(3.106b)

In Eqs. (3.106a) and (3.106b) matrix \( R(i) \) now becomes the Fisher information matrix as compared to the Hesse matrix defined by Eq. (3.105). In the Gauss-Newton method, this matrix will be used in Eq. (3.104), as an approximation of the Hesse matrix in the Newton-Raphson method.

The Gauss-Newton method saves much computation time, since rather than the evaluation of the exact Hesse matrix (3.105), only first order partial derivatives of the negative logarithm of the likelihood function need to be calculated, see Eq. (3.106b).

It is possible to derive an expression for the Fisher information matrix in terms of first order derivatives of the innovations generated by the Kalman filter with respect to the parameters vector \( \theta \), see Eq. (3.106b). Identical expressions result whether the first definition of the Fisher information matrix (based on the second order partial derivatives), see Eq. (3.106a) or the second definition of the Fisher information matrix (based on the first order partial derivatives), see Eq. (3.106b), is used. Below, the expression is derived, starting from the first definition of the Fisher information matrix in (3.106a).

From (3.88), the first order gradient of the likelihood function with respect to the parameters in the parameter vector \( \theta \) can be obtained as:

\[
\frac{\partial}{\partial \theta_i} [L(\theta)] = \sum_{k=1}^{N-1} \left[ \frac{\partial}{\partial \theta_i} [\tilde{y}^T(k+1|k)] \right] \tilde{y}^{-1}(k+1|k) \tilde{y}(k+1|k) + \\
+ \frac{1}{2} \tilde{y}^T(k+1|k) \frac{\partial}{\partial \theta_i} [\tilde{y}^{-1}(k+1|k)] \tilde{y}(k+1|k) + \\
+ \frac{1}{2} \text{det} \left[ \tilde{y}^{-1}(k+1|k) \right] \frac{\partial}{\partial \theta_i} \left\{ \text{det} \left[ \tilde{y}(k+1|k) \right] \right\}, \]  

(3.107)

where \( \theta_i \) is the \( i \)th element of \( \theta \).
Referring to [20], the third term of (3.107) has the following equivalent form:

$$
\text{det} \left[ y_e^{-1}(k+1|k) \right] \frac{\partial}{\partial \theta_i} \left[ \text{det} \left[ y_e(k+1|k) \right] \right] = \text{tr} \left[ y_e^{-1}(k+1|k) \right] \frac{\partial}{\partial \theta_i} \left[ y_e(k+1|k) \right]
$$

(3.108)

and the first order derivatives of the inverted covariance matrix of the prediction errors can be written as:

$$
\frac{\partial}{\partial \theta_i} \left[ y_e^{-1}(k+1|k) \right] = - y_e^{-1}(k+1|k) \frac{\partial}{\partial \theta_i} \left[ y_e(k+1|k) \right] y_e^{-1}(k+1|k) \quad (3.109)
$$

Therefore, it follows now for the first order gradient of the likelihood function that:

$$
\frac{\partial}{\partial \theta_i} [L(\theta)] =
$$

$$
= \sum_{k=0}^{N-1} \left[ \frac{\partial}{\partial \theta_i} \left[ y_e^{-1}(k+1|k) \right] y_e^{-1}(k+1|k) \right] +
$$

$$
- \frac{1}{2} y_e^{-1}(k+1|k) \frac{\partial}{\partial \theta_i} \left[ y_e(k+1|k) \right] y_e^{-1}(k+1|k) +
$$

$$
+ \frac{1}{2} \text{tr} \left[ y_e^{-1}(k+1|k) \right] \frac{\partial}{\partial \theta_i} \left[ y_e(k+1|k) \right].
$$

(3.110)

The second order gradient of the likelihood function with respect to parameters $\theta_i$ and $\theta_j$ of the parameter vector $\theta$ can be obtained from (3.107) as follows:

$$
\frac{\partial^2}{\partial \theta_i \partial \theta_j} [L(\theta)] =
$$
\[
\frac{1}{2} \sum_{k=0}^{N-1} \left[ 2 \left( \frac{\partial^2}{\partial \theta_i \partial \theta_j} \left[ \tilde{y}^T(k+1|k) \right] \right) \tilde{v}_e^{-1}(k+1|k) \tilde{y}(k+1|k) + \right.
\]
\[
\frac{\partial}{\partial \theta_i} \left[ \tilde{y}^T(k+1|k) \right] \frac{\partial}{\partial \theta_j} \left[ \tilde{v}_e^{-1}(k+1|k) \right] \tilde{y}(k+1|k) + \right.
\]
\[
\frac{\partial}{\partial \theta_i} \left[ \tilde{y}^T(k+1|k) \right] \tilde{v}_e^{-1}(k+1|k) \frac{\partial}{\partial \theta_j} \left[ \tilde{y}(k+1|k) \right] + \right.
\]
\[
\left. - \frac{\partial}{\partial \theta_j} \left[ \tilde{y}^T(k+1|k) \right] \tilde{v}_e^{-1}(k+1|k) \frac{\partial}{\partial \theta_i} \left[ \tilde{y}(k+1|k) \right] \tilde{v}_e^{-1}(k+1|k) \tilde{y}(k+1|k) + \right.
\]
\[
- \tilde{y}^T(k+1|k) \tilde{v}_e^{-1}(k+1|k) \frac{\partial^2}{\partial \theta_i \partial \theta_j} \left[ \tilde{v}_e(k+1|k) \right] \tilde{v}_e^{-1}(k+1|k) \tilde{y}(k+1|k) + \right.
\]
\[
+ \frac{\partial}{\partial \theta_j} \left[ \det \left[ \tilde{v}_e^{-1}(k+1|k) \right] \right] \frac{\partial}{\partial \theta_i} \left[ \det \left[ \tilde{v}_e(k+1|k) \right] \right] \right]
\]
\] (3.111)

After some arrangements we obtain:

\[
\frac{\partial^2}{\partial \theta_i \partial \theta_j} [L(\theta)] =
\]
\[
\sum_{k=0}^{N-1} \left[ \frac{\partial}{\partial \theta_i} \left[ \tilde{y}^T(k+1|k) \right] \tilde{v}_e^{-1}(k+1|k) \frac{\partial}{\partial \theta_j} \left[ \tilde{y}(k+1|k) \right] + \right.
\]
\[
\frac{\partial^2}{\partial \theta_i \partial \theta_j} \left[ \tilde{y}^T(k+1|k) \right] \tilde{v}_e^{-1}(k+1|k) + \frac{\partial}{\partial \theta_i} \left[ \tilde{y}^T(k+1|k) \right] \frac{\partial}{\partial \theta_j} \left[ \tilde{v}_e^{-1}(k+1|k) \right] + \right.
\]
\[
\left. + \frac{\partial}{\partial \theta_j} \left[ \tilde{y}^T(k+1|k) \right] \frac{\partial}{\partial \theta_i} \left[ \tilde{v}_e^{-1}(k+1|k) \right] \tilde{y}(k+1|k) + \right.
\]
\[
+ \frac{1}{2} \text{tr} \left[ \frac{\partial}{\partial \theta_j} \left[ \tilde{v}_e^{-1}(k+1|k) \right] \frac{\partial}{\partial \theta_i} \left[ \tilde{v}_e(k+1|k) \right] \right] + \right.
\]
\[ + \mathbf{v}_{\varepsilon}^{-1}(k+1|k) \frac{\partial^2}{\partial \theta_i \partial \theta_j} [\mathbf{v}_{\varepsilon}(k+1|k)] + \]

\[ + \text{tr} \left\{ \left[ \frac{\partial}{\partial \theta_j} [\mathbf{v}_{\varepsilon}^{-1}(k+1|k)] \mathbf{v}_{\varepsilon}(k+1|k) \frac{\partial}{\partial \theta_i} [\mathbf{v}_{\varepsilon}^{-1}(k+1|k)] \right] \right\} + \]

\[ - \frac{1}{2} \mathbf{v}_{\varepsilon}^{-1}(k+1|k) \frac{\partial^2}{\partial \theta_i \partial \theta_j} [\mathbf{v}_{\varepsilon}(k+1|k)] \mathbf{v}_{\varepsilon}^{-1}(k+1|k) \tilde{\mathbf{y}}(k+1|k) \tilde{\mathbf{y}}^T(k+1|k) \right\} \]

(3.112)

In (3.112) the following identities have been used:

\[ \frac{\partial}{\partial \theta_j} [\det \mathbf{v}_{\varepsilon}^{-1}(k+1|k)] \frac{\partial}{\partial \theta_i} [\det \mathbf{v}_{\varepsilon}(k+1|k)] = \]

\[ = \text{tr} \left\{ \frac{\partial}{\partial \theta_j} [\mathbf{v}_{\varepsilon}^{-1}(k+1|k)] \frac{\partial}{\partial \theta_i} [\mathbf{v}_{\varepsilon}(k+1|k)] \right\} + \]

\[ + \text{tr} \left\{ \mathbf{v}_{\varepsilon}^{-1}(k+1|k) \frac{\partial^2}{\partial \theta_i \partial \theta_j} [\mathbf{v}_{\varepsilon}(k+1|k)] \right\} , \]

(3.113)

\[ \frac{\partial}{\partial \theta_j} \left[ \tilde{\mathbf{y}}^T(k+1|k) \mathbf{v}_{\varepsilon}^{-1}(k+1|k) \right] \frac{\partial}{\partial \theta_i} [\mathbf{v}_{\varepsilon}(k+1|k)] \mathbf{v}_{\varepsilon}^{-1}(k+1|k) \tilde{\mathbf{y}}(k+1|k) = \]

\[ = - \left[ \frac{\partial}{\partial \theta_j} \left[ \tilde{\mathbf{y}}^T(k+1|k) \mathbf{v}_{\varepsilon}^{-1}(k+1|k) \right] \right] \tilde{\mathbf{y}}(k+1|k) + \]

\[ + \tilde{\mathbf{y}}^T(k+1|k) \frac{\partial}{\partial \theta_j} [\mathbf{v}_{\varepsilon}^{-1}(k+1|k)] \mathbf{v}_{\varepsilon}(k+1|k) \frac{\partial}{\partial \theta_i} [\mathbf{v}_{\varepsilon}^{-1}(k+1|k)] \tilde{\mathbf{y}}(k+1|k) \right\} , \]

(3.114)

and:

\[ \tilde{\mathbf{y}}^T \mathbf{M} \tilde{\mathbf{y}} = \text{tr} \left( \mathbf{M} \tilde{\mathbf{y}} \tilde{\mathbf{y}}^T \right) , \]

(3.115)
where $M$ is any square matrix and $v$ is any vector with the same dimension as matrix $M$.

In the following analysis, only an approximation of the Fisher information matrix is obtained, since exact evaluation of the expectation of (3.112) is impossible, unless the system is deterministic (Appendix C gives an exact expression of the Fisher information matrix for a deterministic system). The approximation assumes that:

\[
\frac{\partial}{\partial \theta_i} [\hat{y}(k+1|k)] = \frac{\partial}{\partial \theta_i} [y(k+1) - \hat{v}(k+1|k)] = -\frac{\partial}{\partial \theta_i} [\hat{v}(k+1|k)], \tag{3.116}
\]

and:

\[
\frac{\partial^2}{\partial \theta_i \partial \theta_j} [\hat{y}(k+1|k)] = \frac{\partial^2}{\partial \theta_i \partial \theta_j} [y(k+1) - \hat{v}(k+1|k)] = -\frac{\partial^2}{\partial \theta_i \partial \theta_j} [\hat{v}(k+1|k)], \tag{3.117}
\]

are deterministic vectors. Due to this approximation, the expectation of (3.112) may be written as:

\[
E \left[ \frac{\partial^2}{\partial \theta_i \partial \theta_j} [L(\theta)] \right] =
\]

\[
= \sum_{k=0}^{N-1} \left[ \frac{\partial}{\partial \theta_i} \left[ \hat{y}^T(k+1|k) \right] \right] \frac{1}{\sigma_e^2(k+1|k)} \frac{\partial}{\partial \theta_j} [\hat{v}(k+1|k)] +
\]

\[
+ \left[ \frac{\partial^2}{\partial \theta_i \partial \theta_j} \left[ \hat{y}^T(k+1|k) \right] \right] \frac{1}{\sigma_e^2(k+1|k)} \frac{\partial}{\partial \theta_i} \left[ \hat{v}^T(k+1|k) \right] \frac{\partial}{\partial \theta_j} \left[ \hat{v}^{-1}(k+1|k) \right] +
\]

\[
+ \frac{\partial}{\partial \theta_j} \left[ \hat{y}^T(k+1|k) \right] \frac{\partial}{\partial \theta_i} \left[ \hat{v}^{-1}(k+1|k) \right] \frac{\partial}{\partial \theta_i} \left[ \hat{v}^{-1}(k+1|k) \right] \] E \left[ \hat{v}(k+1|k) \right] +
\]

\[
+ \frac{1}{2} \text{tr} \left[ \frac{\partial}{\partial \theta_j} \left[ \hat{v}^{-1}(k+1|k) \right] \right] \frac{\partial}{\partial \theta_i} \left[ \hat{v}^{-1}(k+1|k) \right] +
\]
\[ + \frac{1}{2} \mathbb{V}^{-1}(k+1|k) \frac{\partial^2}{\partial \theta_1 \partial \theta_j} [\mathbb{V}_e(k+1|k)] + \]

\[ + \text{tr} \left\{ \frac{\partial^2}{\partial \theta_j} [\mathbb{V}_e^{-1}(k+1|k)] \mathbb{V}_e(k+1|k) \frac{\partial^2}{\partial \theta_1} [\mathbb{V}_e^{-1}(k+1|k)] \right\} + \]

\[ - \frac{1}{2} \mathbb{V}^{-1}(k+1|k) \frac{\partial^2}{\partial \theta_1 \partial \theta_j} [\mathbb{V}_e(k+1|k)] \mathbb{V}_e^{-1}(k+1|k)] (k+1|k) \mathbb{V}_e^{-1}(k+1|k)] E \{ \bar{y}(k+1,k) \bar{y}^T(k+1|k) \} \} \].

(3.118)

From the properties of the Kalman filter [18], it follows that the state estimates are unbiased, i.e.:

\[ E \{ \bar{y}(k+1|k) \} = 0. \]  

(3.119)

Using (3.86) and (3.119), the \( i, j \)th element of the approximated Fisher information matrix (3.106) may be written as:

\[ R_{i,j} = \sum_{k=0}^{N-1} \left\{ \frac{\partial}{\partial \theta_i} [\bar{y}^T(k+1|k)] \mathbb{V}_e^{-1}(k+1,k) \frac{\partial}{\partial \theta_j} [\bar{y}(k+1|k)] + \right\} + \]

\[ + \frac{1}{2} \text{tr} \left\{ \mathbb{V}_e^{-1}(k+1|k) \frac{\partial}{\partial \theta_j} [\mathbb{V}_e(k+1|k)] \mathbb{V}_e^{-1}(k+1|k) \frac{\partial}{\partial \theta_j} [\mathbb{V}_e(k+1|k)] \right\} + \]

(3.120)

Now the Fisher information matrix \( \mathbb{R} \) only contains the first order derivatives of the state estimator, described by (3.92) to (3.100).

The Gauss-Newton method is probably the most commonly used method for maximum likelihood estimation (see e.g. [15], [16], [19], [26], and [29]). Since \( \mathbb{R} \) is non-negative definite, one can always find an \( \alpha \) such that

\[ L(\theta) \bigg|_{\theta = \theta(i+1)} < L(\theta) \bigg|_{\theta = \theta(i)} \quad \text{.} \]

The method, however, does run into problems when \( \mathbb{R} \) is singular. If \( \mathbb{R} \) is singular, it can not be inverted. If \( \mathbb{R} \)
is nearly singular, the step size $R^{-1}g$ is very large in the near-singular directions. This can be shown as follows.

The Fisher information matrix $R$ can be decomposed as:

$$R = \sum_{j=1}^{m} \lambda_j T_j T_j^T,$$

or:

$$R^{-1} = \sum_{j=1}^{m} \frac{1}{\lambda_j} T_j T_j^T,$$

(3.121)

(3.122)

where $\lambda_j$ and $T_j$ are an eigenvalue of the Fisher information matrix $R$ and the corresponding eigenvector respectively, and $m$ is the number of parameters to be estimated or the dimension of the Fisher information matrix. If the matrix $R$ is nearly singular, at least one eigenvalue is very small as compared to the largest one. From (3.122) it follows, that the algorithm takes a large step in this parameter direction, about which least information is available. This causes the convergence rate of the algorithm to be very poor. It may even lead to divergence. This problem will be discussed in more detail in the next section.

To calculate the gradient of the negative logarithm likelihood function (3.110) as well as of the Fisher information matrix (3.120), the first order partial derivatives of the Kalman filter (3.92) to (3.100) have to be known. These derivatives can be calculated as follows.

Starting from (3.92), the first order partial derivative of the one stage prediction algorithm with respect to the parameter $\theta_i$ can be written from Eq. (3.92) as:

$$\frac{\partial}{\partial \theta_i} \hat{x}(k+1|k) = \left[ \frac{\partial}{\partial \theta_i} (\hat{x}) - K_p \frac{\partial}{\partial \theta_i} (H) \right] \hat{x}(k|k) + \left( \hat{x} - K_p H \right) \frac{\partial}{\partial \theta_i} [\hat{x}(k|k)] +$$

$$+ \left[ \frac{\partial}{\partial \theta_i} (u) - K_p \frac{\partial}{\partial \theta_i} (p) \right] u(k) + \frac{\partial}{\partial \theta_i} (K_p) \left[ y(k) - \hat{y}(k|k) \right].$$

(3.123)
The first order partial derivative of the one stage prediction gain with respect to the parameter \( \theta_i \) can be derived from Eq. (3.94) as:

\[
\frac{\partial}{\partial \theta_i} (\hat{K}) = \frac{\partial}{\partial \theta_i} (\Gamma_w) \frac{\partial}{\partial \theta_i} (V_{ww})^{-1} + \Gamma_w \frac{\partial}{\partial \theta_i} (V_{wv})^{-1} - \Gamma_w V_{wv} V_{v}^{-1} \frac{\partial}{\partial \theta_i} (V_v) V_{v}^{-1} \\
= \frac{\partial}{\partial \theta_i} (\Gamma_w) \frac{\partial}{\partial \theta_i} (V_{ww})^{-1} + \Gamma_w \frac{\partial}{\partial \theta_i} (V_{wv})^{-1} - K_p \frac{\partial}{\partial \theta_i} (V_v) V_{v}^{-1} \ . \quad (3.124)
\]

The first order partial derivative of the measurement update algorithm with respect to the parameter \( \theta_i \) follows from Eq. (3.95) as:

\[
\frac{\partial}{\partial \theta_i} [\hat{x}[k+1|k+1]] = \\
= [\hat{u} - K(k+1) \hat{H}] \frac{\partial}{\partial \theta_i} [\hat{x}(k+1|k)] + \frac{\partial}{\partial \theta_i} [K(k+1)] [\hat{x}(k+1) - \hat{x}(k+1|k)] + \\
- K(k+1) \left[ \frac{\partial}{\partial \theta_i} (\hat{H}) \hat{x}(k+1|k) + \frac{\partial}{\partial \theta_i} (\hat{H}) u(k+1) \right] \ . \quad (3.125)
\]

From (3.96b) or (3.100) the first order partial derivative of the Kalman gain \( K(k+1) \) with respect to the parameter \( \theta_i \) may be derived as:

\[
\frac{\partial}{\partial \theta_i} [K(k+1)] = \left[ \frac{\partial}{\partial \theta_i} [P(k+1|k)] H_T + P(k+1|k) \frac{\partial}{\partial \theta_i} (H_T) \right] + \\
- K(k+1) \frac{\partial}{\partial \theta_i} [V_e(k+1|k)] V_e^{-1}(k+1|k) \ . \quad (3.126)
\]

With (3.91), the first order partial derivative of the covariance matrix of the prediction errors (innovations) with respect to the parameter \( \theta_i \) can be written as:
\[
\frac{\partial}{\partial \theta_1} [\psi(e^{k+1|k})] = \frac{\partial}{\partial \theta_1} [\psi(h)] \frac{\partial}{\partial \theta_1} [\psi(k+1|k)] + \frac{\partial}{\partial \theta_1} [\psi(k+1|k)]^T + \\
+ \frac{\partial}{\partial \theta_1} [\psi(k+1|k)]^T + \frac{\partial}{\partial \theta_1} [\psi(v)] .
\]

(3.127)

The first order partial derivative of the covariance matrix of the prediction errors of the state vector in (3.97), with respect to the parameter \( \theta_1 \), can be written as:

\[
\frac{\partial}{\partial \theta_1} [\psi(k+1|k)] = \\
= \frac{\partial}{\partial \theta_1} [\psi - K_p H] \psi(k|k) \psi - K_p H)^T + \frac{\partial}{\partial \theta_1} [\psi - K_p H] \psi(k|k) \psi - K_p H)^T + \\
+ \frac{\partial}{\partial \theta_1} [\psi - K_p H] \psi(k|k) \psi - K_p H)^T + \frac{\partial}{\partial \theta_1} [\psi - K_p H] \psi(k|k) \psi - K_p H)^T + \\
- \frac{\partial}{\partial \theta_1} [\psi - K_p H] \psi(k|k) \psi - K_p H)^T + \frac{\partial}{\partial \theta_1} [\psi - K_p H] \psi(k|k) \psi - K_p H)^T + \\
- \frac{\partial}{\partial \theta_1} [\psi - K_p H] \psi(k|k) \psi - K_p H)^T + \frac{\partial}{\partial \theta_1} [\psi - K_p H] \psi(k|k) \psi - K_p H)^T + \\
\]

(3.128)

where:

\[
\frac{\partial}{\partial \theta_1} [\psi - K_p H] = \frac{\partial}{\partial \theta_1} [\psi] - \frac{\partial}{\partial \theta_1} [K_p H] - K_p \frac{\partial}{\partial \theta_1} [H] .
\]

(3.129)

Next, the first order partial derivative is derived of the a posteriori covariance matrix (3.99) with respect to the parameter \( \theta_1 \):

\[
\frac{\partial}{\partial \theta_1} [\psi(k+1|k+1)] = \left[ \psi - K(k+1) H \right] \frac{\partial}{\partial \theta_1} [\psi(k+1|k)] + \\
- \left[ \frac{\partial}{\partial \theta_1} [K(k+1)] H + K(k+1) \frac{\partial}{\partial \theta_1} [H] \right] \psi(k+1|k) .
\]

(3.130)
Finally, the first order partial derivative of the prediction errors of the outputs (innovations) with respect to the parameter $\theta_1$ is derived from Eq. (3.83) and (3.93) as:

$$
\frac{\partial}{\partial \theta_1} [\tilde{Y}(k+1|k)] = - \left\{ \frac{\partial}{\partial \theta_1} [H \hat{x}(k+1|k)] + H \frac{\partial}{\partial \theta_1} [\hat{x}(k+1|k)] + \frac{\partial}{\partial \theta_1} (P) u(k+1) \right\}.
$$

(3.131)

In the whole approach described here, the Kalman filter is assumed to be non-stationary. A simplification results, however, if the data collection time is long enough to allow the Kalman filter to become stationary. The negative logarithm of the likelihood function (3.88) can be written in its stationary form as follows:

$$
L(\theta) = \frac{N}{2} \ln [\det (V_e)] + \frac{1}{2} \sum_{k=0}^{N-1} [\tilde{Y}^T(k+1,k) V_e^{-1} \tilde{Y}(k+1,k)],
$$

(3.132)

where the covariance matrix of the prediction errors is now time invariable, and is of the form:

$$
V_e = H P H^T + V_v.
$$

(3.133)

To calculate the stationary covariance matrix of the prediction errors of the states $P$, a steady state matrix Riccati equation must be solved. This equation can be obtained by substituting (3.100) into (3.98) as follows:

$$
P = (\phi - K_p H) \left[ P - P H^T (H P H^T + V_v)^{-1} H P \right] (\phi - K_p H)^T + \Gamma_w V_e \Gamma_w^T - K_v V_v K_v^T
$$

$$
= (\phi - K_p H) \left[ P - P H^T V_e^{-1} H P \right] (\phi - K_p H)^T + \Gamma_w V_w \Gamma_w^T - K_p V_v K_p^T.
$$

(3.134)
In (3.134), the following identity has been used (see [18]):

\[
P = P H^T (H P H^T + V_v)^{-1} H P = P - P H^T V_v^{-1} H P = (P - H^T V_v^{-1} H)^{-1}
\]

(3.135)

Now, Eq. (3.134) can be written in the following form:

\[
P = (\phi - K_p H) (P^{-1} + H^T V_v^{-1} H)^{-1} (\phi - K_p H)^T + \Gamma_w V_v \Gamma_w^T - K_p V_v K_p^T = \rho_P \rho_P^T \]

(3.136)

Next, the steady state Kalman filter gain can be written as:

\[
K = P H^T (H P H^T + V_v)^{-1} = P H^T V_v^{-1}.
\]

(3.137)

The first order partial derivative of the steady state Riccati equation (3.136) with respect to the parameter \( \theta_i \) leads to a matrix Lyapunov equation, see [27]. This equation can be obtained as follows:

\[
\frac{\partial}{\partial \theta_i} \left| P \right| = \frac{\partial}{\partial \theta_i} \left| (\phi - K_p H) (P^{-1} + H^T V_v^{-1} H)^{-1} (\phi - K_p H)^T \right|
\]

\[
= \left[ \frac{\partial}{\partial \theta_i} (\phi - K_p H) (P^{-1} + H^T V_v^{-1} H)^{-1} (\phi - K_p H)^T \right]^T +
\]

\[
- (\phi - K_p H) (P^{-1} + H^T V_v^{-1} H)^{-1} \left| \frac{\partial}{\partial \theta_i} (P) \right| P^{-1} +
\]

\[
+ \frac{\partial}{\partial \theta_i} (H^T V_v^{-1} H) + \left[ \frac{\partial}{\partial \theta_i} (H^T) V_v^{-1} H \right]^T +
\]

\[
- H V_v^{-1} \frac{\partial}{\partial \theta_i} (V_v) V_v^{-1} H \left| (P^{-1} + H^T V_v^{-1} H)^{-1} \right| (\phi - K_p H)^T +
\]
\[ + \frac{\partial}{\partial \theta_1} (\Gamma_w) \begin{bmatrix} \alpha \lambda w \\ \alpha_w \end{bmatrix} T \begin{bmatrix} \alpha \lambda w \\ \alpha_w \end{bmatrix} + \frac{\partial}{\partial \theta_1} (\Gamma_w) \begin{bmatrix} \alpha \lambda w \\ \alpha_w \end{bmatrix} T \begin{bmatrix} \alpha \lambda w \\ \alpha_w \end{bmatrix} T - \frac{\partial}{\partial \theta_1} (\Gamma_w) \begin{bmatrix} \alpha \lambda w \\ \alpha_w \end{bmatrix} T \begin{bmatrix} \alpha \lambda w \\ \alpha_w \end{bmatrix} T + \begin{bmatrix} \alpha \lambda w \\ \alpha_w \end{bmatrix} \frac{\partial}{\partial \theta_1} (\Gamma_w) \begin{bmatrix} \alpha \lambda w \\ \alpha_w \end{bmatrix} T . \quad (3.138) \]

After some arrangements of (3.138), a stationary Lyapunov equation can be derived as:

\[ \psi \frac{\partial}{\partial \theta_1} (p) \psi^T - \frac{\partial}{\partial \theta_1} (p) = a_i + a_i^T + b_i . \quad (3.139) \]

where:

\[ \psi = (\phi - K_p H) (p^{-1} + H^T V^{-1} H)^{-1} p^{-1} \]

\[ = (\phi - K_p H) [U - P H^T (H P H^T + V)^{-1} H] \]

\[ = (\phi - K_p H) (U - P H^T V^{-1} H) \]  
\[ = (\phi - K_p H) (U - P H^T V^{-1} H) \]  
\[ a_i = [(\phi - K_p H) (p - P H^T V^{-1} H P) \frac{\partial}{\partial \theta_1} (H^T) V^{-1} H + \]

\[ - \frac{\partial}{\partial \theta_1} (\phi - K_p H)] (p - P H^T V^{-1} H P) (\phi - K_p H)^T + \]

\[ - \frac{\partial}{\partial \theta_1} (\Gamma_w) V V^T - \frac{\partial}{\partial \theta_1} (K_p V V^T - (\phi - K_p H) (p + \]

\[ b_i = K_p \frac{\partial}{\partial \theta_1} (V) V^T - \Gamma_w \frac{\partial}{\partial \theta_1} (\Gamma_w) V V^T - (\phi - K_p H) (p + \]
\[- p \mu^T v^{-1} H p \) \mu^T v^{-1} \frac{\partial}{\partial \theta_i} (v) v^{-1} H (p +
\]

\[- p \mu^T v^{-1} H p \) (\Phi - k_p p H)^T . \tag{3.142}

The element of the gradient of the negative logarithm of the likelihood function, with respect to the parameter \( \theta_i \) in the stationary case, is of the following form:

\[
\frac{\partial}{\partial \theta_i} [-L(\theta)] = \sum_{k=0}^{N-1} \left\{ \frac{\partial}{\partial \theta_i} \left[ \bar{y}(k+1|k) \right] v^{-1} v(k+1|k) + \right.
\]

\[
- \frac{1}{2} \bar{y}(k+1|k) v^{-1} \frac{\partial}{\partial \theta_i} (v) v^{-1} v(k+1|k) \right\} +
\]

\[
+ \frac{N}{2} tr \left[ v^{-1} \frac{\partial}{\partial \theta_i} (v) \right] , \tag{3.143}
\]

and the \( i, j \) element of the Fisher information matrix in the stationary case is of the form:

\[
R_{i,j} = \sum_{k=0}^{N-1} \left\{ \frac{\partial}{\partial \theta_i} \left[ \bar{y}^T(k+1|k) \right] v^{-1} \frac{\partial}{\partial \theta_j} \left[ \bar{y}(k+1|k) \right] \right\} +
\]

\[
+ \frac{N}{2} tr \left[ v^{-1} \frac{\partial}{\partial \theta_j} (v) v^{-1} \frac{\partial}{\partial \theta_i} (v) \right] . \tag{3.144}
\]

Both the nonstationary and stationary algorithms considered above are, in fact, all based on solving the Riccati and Lyapunov equations (either nonstationary or stationary). In cases where the order of the system equation (3.73) and (3.74) is not too large, there is another approach that may save much more computation time, see [22]. In this approach, the elements of the covariance matrix of the prediction errors of the outputs (innovations) \( v_e \).
the one stage prediction gain \( K_p \) and the Kalman filter gain \( K \) are interpreted as unknown matrices to be estimated. In the total parameter vector \( \theta \) to be estimated in this approach, the elements of the covariance matrices \( V_w, V_v, \) and \( V_wv \) are replaced by the elements of the matrices \( V_e, K_p \) and \( K \). However, the noise covariance matrices \( V_w, V_v, \) and \( V_wv \) can subsequently be reconstructed from the ML estimates of the new parameter vector, see Chapter 4.

The new parameter vector to be estimated is:

\[
\theta = [\theta_a^T \theta_{ve}^T \theta_{kp}^T \theta_k^T]^T ,
\]

where:

- \( \theta_{ve} \) contains the unknown elements of the covariance matrix of the innovations \( V_e \),
- \( \theta_{kp} \) contains the elements of the one stage prediction gain matrix \( K_p \),
- \( \theta_k \) contains the elements in the Kalman gain matrix \( K \), and
- \( \theta_a \) contains the parameters in parameter vectors \( \theta_p, \theta_B, \theta_H, \theta_D \) and \( \theta_E \) defined earlier.

By taking the partial derivative of (3.132) with respect to \( V_e^{-1} \), it follows that the covariance matrix of the innovations can be estimated directly from the corresponding necessary condition for the extreme of the negative logarithm of the likelihood function:

\[
\frac{\partial}{\partial V_e^{-1}} [L(\theta)] = \frac{N}{2} V_e - \frac{1}{2} \sum_{k=0}^{N-1} [\bar{y}(k+1|k) \bar{y}^T(k+1|k)] .
\]

From (3.101) the necessary condition for the extreme is, that the left hand side of (3.146) should be zero, resulting in:

\[
\hat{V}_e = \frac{1}{N} \sum_{k=0}^{N-1} [\bar{y}(k+1|k) \bar{y}^T(k+1|k)] .
\]
The first order gradient of the negative logarithm of the likelihood function with respect to the parameters in $\theta_a$ is:

$$
\frac{\partial}{\partial \theta_a} [L(\theta)] = \sum_{k=0}^{N-1} \frac{\partial}{\partial \theta_a} \left[ \tilde{y}^T(k+1|k) \right] \tilde{v}^{-1}_e \tilde{y}(k+1|k) ,
$$

(3.148)

where:

$$
\frac{\partial}{\partial \theta_a} \left[ \tilde{y}(k+1|k) \right] = - \left[ \frac{\partial}{\partial \theta_a} [H] \right] \hat{x}(k+1|k) + H \frac{\partial}{\partial \theta_a} \left[ \hat{x}(k+1|k) \right] +
$$

$$
+ \frac{\partial}{\partial \theta_a} \left[ D \right] \tilde{y}(k+1) .
$$

(3.149)

The first order gradient with respect to the parameters in $\theta_{kp}$ is:

$$
\frac{\partial}{\partial \theta_{kp}} [L(\theta)] = \sum_{k=0}^{N-1} \frac{\partial}{\partial \theta_{kp}} \left[ \tilde{y}^T(k+1|k) \right] \tilde{v}^{-1}_e \tilde{y}(k+1|k) ,
$$

(3.150)

where:

$$
\frac{\partial}{\partial \theta_{kp}} \left[ \tilde{y}(k+1|k) \right] = - H \frac{\partial}{\partial \theta_{kp}} \left[ \hat{x}(k+1|k) \right] .
$$

(3.151)

Finally, the first order gradient of the negative logarithm of the likelihood function with respect to the parameters in $\theta_k$ is:

$$
\frac{\partial}{\partial \theta_k} [L(\theta)] = \sum_{k=0}^{N-1} \frac{\partial}{\partial \theta_k} \left[ \tilde{y}^T(k+1|k) \right] \tilde{v}^{-1}_e \tilde{y}(k+1|k) ,
$$

(3.152)

where:

$$
\frac{\partial}{\partial \theta_k} \left[ \tilde{y}(k+1|k) \right] = - H \frac{\partial}{\partial \theta_k} \left[ \hat{x}(k+1|k) \right] .
$$

(3.153)
The state estimates and their sensitivities can be written as follows. Note that in the present approach, based on [22], it is not necessary to solve the Riccati and Lyapunov equations.

The one stage prediction algorithm and its sensitivities with respect to the different sets of parameters are given by:

\[
\hat{x}(k+1|k) = \Phi \hat{x}(k+1|k) + \Gamma_u u(k) + K_p [\chi(k) - \hat{\chi}(k|k)], \tag{3.154}
\]

\[
\frac{\partial}{\partial \theta_{a_1}} [\hat{x}(k+1|k)] = [\frac{\partial}{\partial \theta_{a_1}} (\Phi) - K_p \frac{\partial}{\partial \theta_{a_1}} (H)] \hat{x}(k|k) + (\Phi - K_p H) \frac{\partial}{\partial \theta_{a_1}} [\hat{x}(k|k)]
+ \left[ \frac{\partial}{\partial \theta_{a_1}} (\Gamma_u) - K_p \frac{\partial}{\partial \theta_{a_1}} (P) \right] u(k), \tag{3.155}
\]

\[
\frac{\partial}{\partial \theta_{k_p_1}} [\hat{x}(k+1|k)] = (\Phi - K_p H) \frac{\partial}{\partial \theta_{k_p_1}} [\hat{x}(k|k)] + \frac{\partial}{\partial \theta_{k_p_1}} (K_p) [\chi(k) - \hat{\chi}(k|k)], \tag{3.156}
\]

\[
\frac{\partial}{\partial \theta_{k}} [\hat{x}(k+1|k)] = (\Phi - K_p H) \frac{\partial}{\partial \theta_{k}} [\hat{x}(k|k)]. \tag{3.157}
\]

The measurement update of the state estimates and its sensitivities may be written as:

\[
\hat{x}(k+1|k+1) = \hat{x}(k+1|k) + K \left[ \chi(k+1) - \hat{\chi}(k+1|k) \right], \tag{3.158}
\]

\[
\frac{\partial}{\partial \theta_{a_1}} [\hat{x}(k+1|k+1)] = \left[ \frac{\partial}{\partial \theta_{a_1}} (U) - K_p \frac{\partial}{\partial \theta_{a_1}} (H) \right] \hat{x}(k+1|k) + \frac{\partial}{\partial \theta_{a_1}} [\hat{x}(k+1|k)]
+ \left[ \frac{\partial}{\partial \theta_{a_1}} (P) u(k+1) \right], \tag{3.159}
\]

\[
- K_p \left[ \frac{\partial}{\partial \theta_{p_1}} (H) \hat{x}(k+1|k) + \frac{\partial}{\partial \theta_{p_1}} (P) u(k+1) \right].
\]
\[
\frac{\partial}{\partial \theta_k} \hat{x}(k+1|k+1) = (U - K \hat{H}) \frac{\partial}{\partial \theta_k} \hat{x}(k+1|k) \tag{3.160}
\]

\[
\frac{\partial}{\partial \theta_i} \hat{x}(k+1|k+1) = (U - K \hat{H}) \frac{\partial}{\partial \theta_i} \hat{x}(k+1|k) + 
\]

\[
+ \frac{\partial}{\partial \theta_i} (K) [y(k+1) - \hat{x}(k+1|k)] \tag{3.161}
\]

In the present approach, the Fisher information matrix takes the following simple form:

\[
R = \sum_{k=0}^{N-1} \left\{ \frac{\partial}{\partial \theta} [\hat{x}^T(k+1|k)] y^{-1} \frac{\partial}{\partial \theta} [\hat{x}(k+1|k)] \right\} \tag{3.162}
\]

Note, that in (3.162) the parameter vector does not contain the parameters in the parameter vector \(\theta_{ve}\), since the covariance matrix can be derived form (3.147). If the Fisher information matrix with respect to the total parameters including parameters in \(\theta_{ve}\), is required, case should be made of Eq. (3.144).

If the system order is low, the present approach is very efficient and simple, partly due to the fact that explicit solution of the Riccati and Lyapunov equations is avoided. However, if the order of the system is high, the total number of the parameters to be estimated may be too large to allow identification of all the parameters, see Section 3.5 below.

### 3.5 Identifiability of the Flexible Spacecraft Parameters

Most of the problems arising in the actual application of the maximum likelihood parameter estimation techniques discussed in the previous sections, can be classified under the subject heading of "identifiability". Identifiability is related to the degree of excitation of the characteristic modes of the system under investigation and the ability to identify the associated parameters. Identifiability is also related to whether the
individual parameters can be identified or whether they can be identified
only in the form of a smaller number of linear combinations.
Parameter identifiability problems usually become manifest in the form of a
singular or nearly singular Fisher information matrix and difficulties in
obtaining accurate parameter estimates from the measurement data.
Identifiability problems may result from poor input sequences and attempts
to identify too many parameters from a given data set of measurements (see,
e.g., [15] and [17]).
One way to deal with the identifiability problem is to compute the eigen-
values of the Fisher information matrix. All eigenvalues should be real,
since the information matrix is a real symmetrical matrix. In parameter
space the eigenvectors of small eigenvalues correspond to those combinations
of parameters which can hardly be identified (see Eq. (3.121) and (3.122)).
A perfect dependency among the parameters would, strictly speaking, result
in a zero eigenvalue of the information matrix, causing it to be singular.
In practice, however, since round-off and other numerical errors prevent the
information matrix from being exactly singular, all the eigenvalues will be
non-zero, with a spread between the smallest and largest eigenvalues pos-
sibly being many orders of magnitude. In such a case, the information matrix
is ill-conditioned. To deal with this identifiability problem, (3.122) may
still be used but one or more of the smallest eigenvalues should be left
out. First $\mathbf{R}^{-1}$ is decomposed as follows:

\[
\mathbf{R}^{-1} = \sum_{j=1}^{m} \frac{1}{\lambda_j} \mathbf{T}_j \mathbf{T}_j^T
\]

\[
= \sum_{j=1}^{m_1} \frac{1}{\lambda_j} \mathbf{T}_j \mathbf{T}_j^T + \sum_{j=m_1+1}^{m_1+m_2} \frac{1}{\lambda_j} \mathbf{T}_j \mathbf{T}_j^T,
\]

(3.163)

where $m_1$ is the number of eigenvalues to be retained, $m_2$ is the number of
eigenvalues which will be left out and $m_1 + m_2 = m$ equals the total number
of the eigenvalues of the Fisher information matrix.
Next, a pseudo inverse of the information matrix $\mathbf{R}$ is computed from only the
first term of Eq. (3.163), i.e.:
Eq. (3.164) indicates that only m₁ parameters can be identified. The remaining m₂ parameters should, therefore, be given some a priori values.

The number of unknown parameters in a flexible spacecraft may be very large, if no a priori information is available, concerning system model parameter values. This may lead to an unmanageable number of parameters, which have to be determined from only a limited number of measurements. Consequently, the Fisher information matrix may hardly be expected to be well-conditioned. As discussed in Chapter 2, the present study is based on the use of quantitative a priori knowledge from finite element analysis, to reduce the number of parameters. As shown in Chapter 4, this results in a manageable number of unknown parameters, which can readily be handled by the maximum likelihood estimation procedure described above.

3.6 Properties of the Maximum Likelihood Estimator

The ML estimates of unknown parameters have the following properties as reviewed and discussed in Ref. [10].

1) They are consistant, i.e.:

\[
\lim_{N \to \infty} P\{|\hat{\theta} - \theta| \leq \varepsilon\} = 1 ,
\]  

where \( P\{ \} \) denotes the probability and \( \theta \) is the true parameter vector.

2) They are asymptotically unbiased, i.e.:

\[
\lim_{N \to \infty} E\{\hat{\theta}\} = \theta .
\]

3) They are asymptotically efficient with:

\[
\lim_{N \to \infty} E\{(\hat{\theta} - \theta)(\hat{\theta} - \theta)^T\} = - E\left\{\frac{\partial}{\partial \theta} \left[ \ln L(\theta) \right] \right\}^{-1} .
\]
The inverse of the Fisher information matrix provides a lower bound on the covariance of the errors in the estimated parameters. This lower bound is the so-called Cramer-Rao lower bound. This means that an ML estimate is the best of all conceivable estimates for large sample sizes (see e.g. [15]).

3.7 Sensitivity Matrix Calculations for the Reduced Order System Model

Section 3.3 makes extensive use of sensitivity matrices. This section discusses the calculation of the sensitivity matrices in more detail.

From Eqs. (3.68), (3.69) and (3.70), it follows that the sensitivity matrices of transition and input distribution matrices with respect to a parameter $\theta_i$ may be derived as follows.

\[
\frac{\partial}{\partial \theta_i}(\Phi) = \sum_{n=1}^{\infty} \frac{1}{n!} \frac{\partial}{\partial \theta_i} \left[ (F^{-1}_{\theta_a} \ A_{\theta_a})^n \ T^n \right]
\]

\[
= \sum_{n=1}^{\infty} \left\{ \frac{1}{n!} \sum_{q=1}^{n} \left\{ (F^{-1}_{\theta_a} \ A_{\theta_a})^{q-1} \ F^{-1}_{\theta_a} \ \left[ \frac{\partial}{\partial \theta_i} (A_{\theta_a}) \right] + \sum_{q=1}^{n-q} \left[ (F^{-1}_{\theta_a} \ A_{\theta_a})^{n-q} \right] \right\} T^n \right\}, \tag{3.168}
\]

\[
\frac{\partial}{\partial \theta_i}(\Gamma_i) = \left\{ \sum_{n=1}^{\infty} \frac{1}{(n+1)!} \right\} \left\{ \sum_{q=1}^{n} \left\{ (F^{-1}_{\theta_a} \ A_{\theta_a})^{q-1} \ F^{-1}_{\theta_a} \ \left[ \frac{\partial}{\partial \theta_i} (A_{\theta_a}) \right] + \sum_{q=1}^{n-q} \left[ (F^{-1}_{\theta_a} \ A_{\theta_a})^{n-q} \right] \right\} T^n \right\} \right\} + \sum_{n=1}^{\infty} \frac{1}{(n+1)!} \left\{ F^{-1}_{\theta_a} \ A_{\theta_a} \right\} T^n \right\} \right\} \right\} \right\}.
\]

\[
\frac{\partial}{\partial \theta_i}(\Gamma_i) = \left\{ \sum_{n=1}^{\infty} \frac{1}{n!} \left\{ (F^{-1}_{\theta_a} \ A_{\theta_a})^n \ T^n \right\} \right\} T^{-1}_{\theta_a} \left[ \frac{\partial}{\partial \theta_i} (B_{\theta_a}) \right] + \sum_{n=1}^{\infty} \frac{1}{n!} \left\{ (F^{-1}_{\theta_a} \ A_{\theta_a})^n \ T^n \right\} \right\} \right\} + \sum_{n=1}^{\infty} \frac{1}{n!} \left\{ (F^{-1}_{\theta_a} \ A_{\theta_a})^n \ T^n \right\} \right\} \right\} \right\}.
\]

\[
\frac{\partial}{\partial \theta_i}(G_i) = \left\{ \sum_{n=1}^{\infty} \frac{1}{n!} \left\{ (F^{-1}_{\theta_a} \ A_{\theta_a})^n \ T^n \right\} \right\} + \sum_{n=1}^{\infty} \frac{1}{n!} \left\{ (F^{-1}_{\theta_a} \ A_{\theta_a})^n \ T^n \right\} \right\} \right\}.
\]

\[
\tag{3.169}
\]
\[
\frac{3}{\theta_1}(\Gamma_w) = \left\{ \sum_{n=1}^{\infty} \left\{ \frac{1}{(n+1)!} \right\} \left\{ \sum_{q=1}^{n} \left( F_a^{-1} A_a \right)^q F_a^{-1} \right\} \frac{3}{\theta_1}(A_a) +
\right.
\left.
- \frac{3}{\theta_1}(F_a) F_a^{-1} A_a \left( F_a^{-1} A_a \right)^{n-q} T^n \right\} T F_a^{-1} G_a +
\right.
\left.
+ \left[ U + \sum_{n=1}^{\infty} \left\{ \frac{1}{(n+1)!} \right\} \left( F_a^{-1} A_a \right)^n T^n \right\} T F_a^{-1} \frac{3}{\theta_1}(F_a) F_a^{-1} G_a .
\right]
\]

(3.170)

The sensitivity matrices of the covariance matrices \( V = W \), \( V = V \) and \( V = W V \) with respect to \( \theta_1 \) can be evaluated from (3.77), (3.78) and (3.79) as follows.

\[
\frac{3}{\theta_1}(V_i = W) = \begin{bmatrix}
\frac{3}{\theta_1}(V_i = W1) & 0_{m_m + m_z} \\
0_{m_z + m_m} & \frac{3}{\theta_1}(V_i = W2)
\end{bmatrix},
\]

(3.171)

\[
\frac{3}{\theta_1}(V_i = V) = \frac{3}{\theta_1}(C_1 F_r^{-1} G_1) V_1 (C_1 F_r^{-1} G_1)^T +
\]

\[
\left[ \frac{3}{\theta_1}(C_1 F_r^{-1} G_1) V_1 (C_1 F_r^{-1} G_1)^T \right]^T +
\]

\[
+ C_1 F_r^{-1} G_1 \frac{3}{\theta_1}(V_i = W1) (C_1 F_r^{-1} G_1) + \frac{3}{\theta_1}(V_i = V1) ,
\]

(3.172)

where:

\[
\frac{3}{\theta_1}(C_1 F_r^{-1} G_1) = \frac{3}{\theta_1}(C_1 F_r^{-1} G_1) F_r^{-1} G_1 - C_1 F_r^{-1} \frac{3}{\theta_1}(F_r^{-1}) F_r^{-1} G_1 ,
\]

(3.173)

and:
\[ \frac{\partial}{\partial \theta} (v_{w1}) = \frac{\partial}{\partial \theta} (v_{w1}) (C_1 E_r^{-1} G_1)^T + v_{w1} \frac{\partial}{\partial \theta} (C_1 E_r^{-1} G_1)^T. \] (3.174)

The sensitivity matrices derived above contain some other sensitivity matrices i.e.: \( \frac{\partial}{\partial \theta} (F_a) \), \( \frac{\partial}{\partial \theta} (A_a) \) and \( \frac{\partial}{\partial \theta} (B_a) \). These matrices are now evaluated as well. From Subsection 3.2.3, it follows that:

\[
A_a = \begin{bmatrix}
  mU_3 & 0 & W_c & 0 & 0 \\
  0 & 1 & A_c & 0 & 0 \\
  W_c^T & A_c^T & M_c & 0 & 0 \\
  0 & 0 & m_r (6 + m_r) & U & 0 \\
  0 & 0 & m_z x 3 & 0 & U \\
\end{bmatrix}
\]

\[
F_a = \begin{bmatrix}
  F_r & 0 & 0 \\
  0 & 0 & 0 \\
  0 & 0 & 0 \\
\end{bmatrix}
\]

\[
A_a = \begin{bmatrix}
  m_r x 3 & -A_c w = o & 0 & 0 & 0 \\
  0 & 3 & 0 & 0 & 0 \\
  0 & -I_w = w = o & 0 & 0 & 0 \\
  0 & 0 & 0 & 0 & 0 \\
  0 & 0 & 0 & 0 & 0 \\
\end{bmatrix}
\]

\[
F_r = \begin{bmatrix}
  F_r & 0 & 0 \\
  0 & 0 & 0 \\
  0 & 0 & 0 \\
\end{bmatrix}
\]
\[
\begin{align*}
M_c &= M_m - K_m K_s^{-1} M_s^T - (K_m K_s^{-1} M_m^T)^T + K_m K_s^{-1} M_m K_s^{-1} K_m^T, \\
K_c &= K_m - K_m K_s^{-1} K_m^T, \\
D_c &= D_m - K_m K_s^{-1} D_m^T - (K_m K_s^{-1} D_m^T)^T + K_m K_s^{-1} D_m K_s^{-1} K_m^T \\
&= \alpha M_c + \beta K_c, \\
W_c &= W_m - W_s K_s^{-1} K_m^T, \\
A_c &= A_m - A_s K_s^{-1} K_m^T.
\end{align*}
\]
In (3.183) the original coupling matrix \( A \) (see Eq. (2.62)) is decomposed into \( A_m \) and \( A_s \), a master and a slave coupling matrix respectively from Eq. (3.8), i.e.:

\[
[A_m ; A_s] = \left( \sum_{i=1}^{N} T_{i1}^T A_{i1} \ p_i \right) T_{se}^T,
\]

(3.184)

in which \( T_{se} \) is an elementary transformation matrix defined by Eq. (3.3):

\[
d = T_{se}^T \begin{bmatrix} d_m \\ \vdots \\ d_s \end{bmatrix}.
\]

(3.185)

In (3.185) \( d_m \) denotes the master displacements to be retained as the reduced displacements \( d_r \) of the flexible substructures, i.e.:

\[
d_r = d_m.
\]

(3.186)

A similar decomposition of the original coupling matrix \( W \) (2.63) can be written from Eq. (3.7) as:

\[
[W_m ; W_s] = \left( \sum_{i=1}^{N} T_{i1}^T W_{i1} \ p_i \right) T_{se}^T.
\]

(3.187)

The sensitivities of these matrices as well as those of the observation matrix \( H \) in Eq. (3.55) and the feedforward matrix \( D \) in Eq. (3.56), with respect to the parameter \( \theta_i \), can now be derived as follows:
\[
\frac{\partial}{\partial \theta_i} (\bar{F}_a) =
\begin{bmatrix}
\frac{\partial}{\partial \theta_i} (m_i) y_3 & \frac{\partial}{\partial \theta_i} (q) & \frac{\partial}{\partial \theta_i} (\bar{w}_c) & 0_{3 \times (6 + m_r + m_z)} \\
\frac{\partial}{\partial \theta_i} (\bar{q}^T) & \frac{\partial}{\partial \theta_i} (\bar{r}) & \frac{\partial}{\partial \theta_i} (\bar{A}_c) & 0_{3 \times (6 + m_r + m_z)} \\
\frac{\partial}{\partial \theta_i} (\bar{w}_c^T) & \frac{\partial}{\partial \theta_i} (\bar{A}_c^T) & \frac{\partial}{\partial \theta_i} (\bar{m}_c) & 0_{m_r \times (6 + m_r + m_z)} \\
0_{(6 + m_r + m_z) \times 3} & 0_{(6 + m_r + m_z) \times 3} & 0_{(6 + m_r + m_z) \times m_r} & 0_{(6 + m_r + m_z)}
\end{bmatrix}
\]

\[
\frac{\partial}{\partial \theta_i} (\bar{r}) = \begin{bmatrix}
0_{(12 + 2m_r) \times m_z} \\
0_{m_z \times (12 + 2m_r)} & 0_{m_z}
\end{bmatrix}
\]

\[(3.188)\]

\[
\frac{\partial}{\partial \theta_i} (\bar{A}_a) =
\begin{bmatrix}
0_3 & -\frac{\partial}{\partial \theta_i} (Q_{\omega}) & 0_{3 \times m_r} & 0_3 & 0_{3 \times m_r} & 0_{3 \times m_z} \\
0_3 & -\frac{\partial}{\partial \theta_i} (I_{\omega}) & 0_{3 \times m_r} & 0_3 & 0_{3 \times m_r} & 0_{3 \times m_z} \\
0_{m_r \times 3} & -\frac{\partial}{\partial \theta_i} (A_{\omega}^T) & 0_{m_r \times 3} & 0_{m_r \times 3} & -\frac{\partial}{\partial \theta_i} (K_c) & 0_{m_r \times m_z} \\
0_3 & 0_3 & 0_{3 \times m_r} & 0_3 & 0_{3 \times m_r} & 0_{3 \times m_z} \\
0_3 & 0_3 & 0_{3 \times m_r} & 0_3 & 0_{3 \times m_r} & 0_{3 \times m_z} \\
0_{m_r \times 3} & 0_{m_r \times 3} & 0_{m_r} & 0_{m_r \times 3} & 0_{m_r} & 0_{m_r \times m_z} \\
0_{m_z \times 3} & 0_{m_z \times 3} & 0_{m_z \times m_r} & 0_{m_z \times 3} & 0_{m_z \times m_r} & \frac{\partial}{\partial \theta_i} (\bar{e})
\end{bmatrix}
\]

\[(3.189)\]
\[ \frac{\partial}{\partial \theta_1} (B_{r}) = \begin{bmatrix} \frac{\partial}{\partial \theta_1} (B_{r})^0_{\mu} \\ \vdots \\ 0_{m_z \times m_u} \end{bmatrix} , \]  

(3.190)

\[ \frac{\partial}{\partial \theta_1} (H) = \left[ \frac{\partial}{\partial \theta_1} (G_1) \right] F_r^{-1} A_r - C_1 F_r^{-1} \left[ \frac{\partial}{\partial \theta_1} (F_r) \right] F_r^{-1} A_r + \frac{\partial}{\partial \theta_1} (A_r) + \frac{\partial}{\partial \theta_1} (G_2) \left[ \frac{\partial}{\partial \theta_1} (F_r) - C_1 F_r^{-1} \frac{\partial}{\partial \theta_1} (F_r) \right] F_r^{-1} G_r \right) . \]  

(3.191)

\[ \frac{\partial}{\partial \theta_1} (D) = \left[ \frac{\partial}{\partial \theta_1} (G_1) \right] F_r^{-1} B_r - C_1 F_r^{-1} \left[ \frac{\partial}{\partial \theta_1} (F_r) \right] F_r^{-1} B_r + \frac{\partial}{\partial \theta_1} (B_r) \right) . \]  

(3.192)

In the expressions of (3.188) and (3.189), the sensitivities of the submatrices can be derived as follows:

\[ \frac{\partial}{\partial \theta_1} (m) = \frac{\partial}{\partial \theta_1} (m_b) + \frac{\partial}{\partial \theta_1} (m_a) = \frac{\partial}{\partial \theta_1} (m_b) + \sum_{j=1}^{N} \left[ \frac{\partial}{\partial \theta_1} (m_j) \right] , \]  

(3.193)

\[ \frac{\partial}{\partial \theta_1} (I) = \frac{\partial}{\partial \theta_1} (I_b) + \frac{\partial}{\partial \theta_1} (I_a) = \]  

\[ = \frac{\partial}{\partial \theta_1} (I_b) + \sum_{j=1}^{N} \left[ T_j^T \frac{\partial}{\partial \theta_1} (I_j) T_j + \frac{\partial}{\partial \theta_1} (T_j^T) I_j T_j \right] , \]  

(3.194)

\[ \frac{\partial}{\partial \theta_1} (Q) = \frac{\partial}{\partial \theta_1} (Q_b) + \frac{\partial}{\partial \theta_1} (Q_a) = \]  

\[ = \frac{\partial}{\partial \theta_1} (Q_b) + \sum_{j=1}^{N} \left[ T_j^T \frac{\partial}{\partial \theta_1} (Q_j) T_j + \frac{\partial}{\partial \theta_1} (T_j^T) Q_j T_j \right] , \]  

(3.195)
\[
+ \left[ \frac{\partial}{\partial \theta_1} (T_j^T \ C_j \ T_j) \right]_T,
\]

\[
\frac{3}{\partial \theta_1} (A_c) = -\frac{3}{\partial \theta_1} (A_m) - \frac{3}{\partial \theta_1} (A_s) \ K_{ss}^{-1} K_{ms}^{-1} + A_s \ K_{ss}^{-1} \left[ \frac{3}{\partial \theta_1} (K_{ss}) \ K_{ss}^{-1} K_{ms}^{-1} \left( K_{ms}^T \right) \right] \]

\[
+ \left[ \frac{3}{\partial \theta_1} (K_{ms}^T) \right].
\]

\[
\frac{3}{\partial \theta_1} (A_m) = \frac{3}{\partial \theta_1} (A_s) \ K_{ss}^{-1} K_{ms}^{-1} + A_s \ K_{ss}^{-1} \left[ \frac{3}{\partial \theta_1} (K_{ss}) \ K_{ss}^{-1} K_{ms}^{-1} \left( K_{ms}^T \right) \right] \]

\[
\frac{3}{\partial \theta_1} (A_s) = \frac{3}{\partial \theta_1} (A_m) \ K_{ss}^{-1} K_{ms}^{-1} + A_s \ K_{ss}^{-1} \left[ \frac{3}{\partial \theta_1} (K_{ss}) \ K_{ss}^{-1} K_{ms}^{-1} \left( K_{ms}^T \right) \right] \]

\[
\frac{3}{\partial \theta_1} (w_c) = \frac{3}{\partial \theta_1} (w_m) - \frac{3}{\partial \theta_1} (w_s) \ K_{ss}^{-1} K_{ms}^{-1} + w_s \ K_{ss}^{-1} \left[ \frac{3}{\partial \theta_1} (K_{ss}) \ K_{ss}^{-1} K_{ms}^{-1} \left( K_{ms}^T \right) \right] \]

\[
- \left[ \frac{3}{\partial \theta_1} (K_{ms}^T) \right].
\]

\[
\frac{3}{\partial \theta_1} (w_m) = \frac{3}{\partial \theta_1} (w_s) \ K_{ss}^{-1} K_{ms}^{-1} + w_s \ K_{ss}^{-1} \left[ \frac{3}{\partial \theta_1} (K_{ss}) \ K_{ss}^{-1} K_{ms}^{-1} \left( K_{ms}^T \right) \right] \]

\[
\frac{3}{\partial \theta_1} (w_s) = \frac{3}{\partial \theta_1} (w_m) \ K_{ss}^{-1} K_{ms}^{-1} + w_s \ K_{ss}^{-1} \left[ \frac{3}{\partial \theta_1} (K_{ss}) \ K_{ss}^{-1} K_{ms}^{-1} \left( K_{ms}^T \right) \right] \]

\[
\frac{3}{\partial \theta_1} (m_c) = \frac{3}{\partial \theta_1} (m_{mm}) - \frac{3}{\partial \theta_1} (K_{ms} \ K_{ss}^{-1} M_{ms}^T) - \frac{3}{\partial \theta_1} (K_{ms} \ K_{ss}^{-1} M_{ms})^T \]

\[
+ \frac{3}{\partial \theta_1} (K_{ms} \ K_{ss}^{-1} M_{ms} \ K_{ms}^{-1} K_{ss}^T). \]

In Eq. (3.200),

\[
\frac{3}{\partial \theta_1} (K_{ms} \ K_{ss}^{-1} M_{ms}^{-1}) = \frac{3}{\partial \theta_1} (K_{ms}) \ K_{ss}^{-1} M_{ms}^T - K_{ms} \ K_{ss}^{-1} \frac{3}{\partial \theta_1} (K_{ss}) \ K_{ss}^{-1} + \]

\[
+ K_{ms} \ K_{ss}^{-1} \frac{3}{\partial \theta_1} (M_{ms}^T). \]
\[ \frac{\partial}{\partial \theta_i} (K_{ms}^{-1} M_{ss} K_{ss}^{-1} K_{ms}^T) = \]

\[ = \frac{\partial}{\partial \theta_i} (K_{ms}) K_{ss}^{-1} M_{ss} K_{ss}^{-1} K_{ms}^T + \left[ \frac{\partial}{\partial \theta_i} (K_{ms}) K_{ss}^{-1} M_{ss} K_{ss}^{-1} K_{ms}^T \right]^T + \]

\[ - K_{ms} K_{ss}^{-1} \frac{\partial}{\partial \theta_i} (K_{ss}) K_{ss}^{-1} M_{ss} K_{ss}^{-1} K_{ms}^T + K_{ms} K_{ss}^{-1} \frac{\partial}{\partial \theta_i} (M_{ss}) K_{ss}^{-1} K_{ms}^T + \]

\[ - [K_{ms} K_{ss}^{-1} \frac{\partial}{\partial \theta_i} (K_{ss}) K_{ss}^{-1} M_{ss} K_{ss}^{-1} K_{ms}^T]_T, \]  

(3.202)

\[
\begin{bmatrix}
\frac{\partial}{\partial \theta_i} (M_{mm}) \\
\frac{\partial}{\partial \theta_i} (M_{ms}) \\
\frac{\partial}{\partial \theta_i} (M_{ms}^T) \\
\frac{\partial}{\partial \theta_i} (M_{ss})
\end{bmatrix}
= T_e \left[ \frac{\partial}{\partial \theta_i} \left( \sum_{j=1}^N p_j \frac{\partial}{\partial \theta_i} (M_j) p_j \right) T_e \right]^T, \]  

(3.203)

\[ \frac{\partial}{\partial \theta_i} (K_c) = \frac{\partial}{\partial \theta_i} (K_{mm}) - \frac{\partial}{\partial \theta_i} (K_{ms}) K_{ss}^{-1} K_{ms}^T - \left[ \frac{\partial}{\partial \theta_i} (K_{ms}) K_{ss}^{-1} K_{ms}^T \right]^T + \]

\[ + K_{ms} K_{ss}^{-1} \frac{\partial}{\partial \theta_i} (K_{ss}) K_{ss}^{-1} K_{ms}^T, \]

(3.204)

\[
\begin{bmatrix}
\frac{\partial}{\partial \theta_i} (K_{mm}) \\
\frac{\partial}{\partial \theta_i} (K_{ms}) \\
\frac{\partial}{\partial \theta_i} (K_{ms}^T) \\
\frac{\partial}{\partial \theta_i} (K_{ss})
\end{bmatrix}
= T_e \left[ \frac{\partial}{\partial \theta_i} \left( \sum_{j=1}^N p_j \frac{\partial}{\partial \theta_i} (K_j) p_j \right) T_e \right]^T, \]  

(3.205)

and:

\[ \frac{\partial}{\partial \theta_i} (D_c) = \frac{\partial}{\partial \theta_i} (\alpha) M_c + \alpha \frac{\partial}{\partial \theta_i} (M_c) + \frac{\partial}{\partial \theta_i} (\beta) K_c + \beta \frac{\partial}{\partial \theta_i} (K_c). \]  

(3.206)
It may be deduced from the above expressions, that the original system is of very high order and the calculation of the sensitivity matrices may be very time consuming. However, considerable simplifications result, if it is assumed that a number of parameters can be measured directly on the ground. This applies in particular to the dimensions of the flexible appendages, the elasticity properties of the materials and the offset angles of the flexible substructures. In that case the remaining sensitivity matrices are:

\[
\frac{\partial}{\partial \theta_i}(m) = \frac{\partial}{\partial \theta_i}(m_b),
\]

\[
\frac{\partial}{\partial \theta_i}(I) = \frac{\partial}{\partial \theta_i}(I_b),
\]

\[
\frac{\partial}{\partial \theta_i}(P_c) = \frac{\partial}{\partial \theta_i}(\alpha) M_c + \frac{\partial}{\partial \theta_i}(\beta) K_c,
\]

\[
\frac{\partial}{\partial \theta_i}(M_c) = 0_{m_r},
\]

\[
\frac{\partial}{\partial \theta_i}(K_c) = 0_{m_r},
\]

\[
\frac{\partial}{\partial \theta_i}(A_c) = 0_{3 \times m_r},
\]

\[
\frac{\partial}{\partial \theta_i}(W_c) = 0_{3 \times m_r},
\]

\[
\frac{\partial}{\partial \theta_i}(Q) = \frac{\partial}{\partial \theta_i}(Q_b).
\]

The sensitivity of the matrix $F_a$ with respect to the parameter $\theta_i$ can now be simplified as follows:
\[
\frac{\partial}{\partial \Theta_1}(F_a) = \begin{bmatrix}
\frac{\partial}{\partial \Theta_1}(m_b) & \frac{\partial}{\partial \Theta_1}(Q_b) & 0 & 3x(6+2m_r+m_z) \\
\frac{\partial}{\partial \Theta_1}(Q^T_c) & \frac{\partial}{\partial \Theta_1}(Q =_b) & 0 & 3x(6+2m_r+m_z) \\
0 & 0 & (6+2m_r+m_z)x3 & 0 \\
0 & 0 & 0 & 6+2m_r+m_z
\end{bmatrix},
\]

(3.215)

and:

\[
\frac{\partial}{\partial \Theta_1}(A_a) = \begin{bmatrix}
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
\frac{\partial}{\partial \Theta_1}(m_r) & \frac{\partial}{\partial \Theta_1}(Q =_o) & 0 & 3x(6+2m_r+m_z) \\
\frac{\partial}{\partial \Theta_1}(A^T_c) & \frac{\partial}{\partial \Theta_1}(Q =_o) & 0 & 3x(6+2m_r+m_z) \\
0 & 0 & (6+2m_r)x3 & 0 \\
0 & 0 & 0 & 6+2m_r+m_z \\
0 & 0 & 0 & 0 \\
0 & 0 & (6+2m_r)x3 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0
\end{bmatrix},
\]

(3.216)

It may be concluded, that the calculation of the sensitivity matrices is much simpler than if all parameters would have to be estimated. The few remaining sensitivities of the system matrices \( F_a \) and \( A_a \) are those with respect to the mass of the rigid body \( m_b \), the inertia matrix of the rigid body \( I =_b \), the coupling matrix of the undeformed spacecraft \( Q =_b \), the damping coefficients \( \alpha \) and \( \beta \), and the parameters in the matrix \( E \).

References


3.22 Milder, J.A. Private Communications.


3.45 Chu, Q.P. "Maximum Likelihood Parameter Estimation Using the Kalman Filter Described by a One Stage Prediction Representation", Technical Report, Delft University of Technology, Delft, The Netherlands, to be Published.
Fig. 4.2 European Space Station (COLUMBUS).
Fig. 4.3 Earth-Observation Platform.
CHAPTER 4

SIMULATION EXPERIMENTS AND RESULTS
OF ML FLEXIBLE SPACECRAFT PARAMETER IDENTIFICATION

4.1 Introduction

The method parameter identification of a flexible spacecraft, as discussed in the previous chapters, is based on three main steps to obtain the best estimation of the parameters of the spacecraft from the measurement data. In Chapter 2, the finite element method is used to generate a dynamical model of the flexible spacecraft. This results in some uncertainties, i.e. Young's modulus, Poisson ratios, mass density, dimensions of elements, etc., in the model matrices (inertia, mass, damping, stiffness and coupling matrices, etc.), which are treated as unknown parameters to be estimated. The estimation of all the elements of the system matrices is then reduced to the estimation of only a reasonable number of parameters to be handled by parameter estimation algorithms, see Section 2.6.

In Chapter 3, the originally high order system model based on a large number of degrees of freedom, is reduced to a model of lower order. Some additional disturbances, which account for the modelling errors resulting from model order reduction, are then added to the reduced order model. In principle, these disturbances are also unknown as discussed in Chapter 3. To provide the possibility to estimate these unknown disturbances, a model of the unknown disturbances has to be assumed first. This model is assumed to consist of white noises process with Gaussian distribution, or treated as a more general form which has been described in Subsection 3.2.3. In both cases, the parameter identification procedure results in the estimation of the parameters of the covariance matrix of white noise process (the model given in 3.2.3 is also excited by white noises). The reduced model with additional disturbances is then used in the third step of the whole estimation process. In this step, the maximum likelihood parameter estimation obtains the best estimates of parameters of the reduced order model to fit the original actual spacecraft. The negative logarithm of the likelihood function indicates the performance index of this fit. The inverted Fisher information matrix
provides the lower bound of the covariance matrix of the errors of the estimated parameters.

In the present chapter, the maximum likelihood parameter estimation algorithms as described in Chapter 3, are applied to some experiments on the parameter estimation of flexible spacecraft. The purpose of these applications is the verification of the applicability of maximum likelihood estimation to flexible spacecraft parameter estimation. The experiments require measurement data from flexible spacecraft.

As no measurement data from a real spacecraft are available, the data are obtained from a simulated hypothetical satellite.

In Section 4.2, a summary of the computer programs for the maximum likelihood parameter estimation of a linear stochastic systems are given. To verify the computer programs, some simple test examples are discussed in Section 4.3.

To generate the measurement data, a high order dynamical model is used to simulate the actual dynamical behaviour of the flexible spacecraft. The time responses are obtained by numerical integration and then corrupted by simulated measurement noises, to yield simulated measured time responses of the assumed flexible spacecraft. These simulation experiments are described in Section 4.4.

To start the maximum likelihood estimation of the flexible spacecraft parameters, initial values of the unknown parameters have to be estimated. Good estimates of the initial parameters can yield good results of ML parameter estimates, as discussed in Chapter 3. The determination of initial values of the parameters is discussed in Section 4.5.

The results of the ML estimation of flexible spacecraft unknown parameters are discussed in Section 4.6. In this section two types of disturbances (white noise process and colored noise) accounting for the modelling errors due to the order reduction, are discussed. The estimated parameters and the estimated system states are given and discussed for the two different types of disturbances.
4.2 A Summary of the Computer Programs for ML Estimation of the Parameters in Linear System with Correlated Process and Measurement

A block diagram of the computer program used to obtain ML estimates of the unknown parameters of the flexible spacecraft model, is given in Fig. 4.1. The algorithms have been discussed in Sections 3.3 to 3.7.

The measured outputs of the spacecraft are compared with the predicted outputs of the spacecraft model using the lower order and estimated parameters. The difference between the time responses is the predicted error (innovation), defined by (3.83). The Gauss-Newton optimization method (Section 3.4) is used to find the parameter values that optimize the negative logarithm of the likelihood function, described by Eq. (3.88) for the nonstationary case or Eq. (3.132), for the stationary case. Each iteration of this algorithm provides new estimates of the unknown parameters on the basis of the prediction error. These new estimates of the parameters are then used to update the model of the flexible spacecraft, providing a new prediction and, therefore, a new prediction error. The updating of the estimated flexible spacecraft model continues iteratively, until a convergence criterion is satisfied. The estimates resulting from this procedure are the maximum likelihood estimates.

According to the different algorithms defined in Chapter 3, different computer programs should be specified. Fig. 4.2 shows the characteristics of the different program options.

Flow charts for each of these algorithms of stochastic system ML parameter estimation are given in Figs. 4.3 to 4.5.

Computer programs for the above algorithms for maximum likelihood parameter estimation, as applied to general linear stochastic system, have been written in both FORTRAN-77* and ALOOIL-60. They are available in the Subject Group of Stability and Control of the Faculty of Aerospace Engineering, TU-Delft. The programs can be applied to general linear system with correlated process and measurement noises.
Fig. 4.1 Parameter estimation using the maximum likelihood concept.

Algorithm Selection

| Nonsteady state ML estimation algorithm (Fig. 4.3) | Steady state estimation algorithm, solving the Ric. and Lya. equations (Fig. 4.4) | Steady state estimation algorithm, without solving the Ric. and Lya. equations (Fig. 4.5) |

Fig. 4.2 Different algorithms for ML estimation.
### Program Flow Chart of the Nonsteady State ML Estimation Algorithm

1. **read initial parameter values, input and output measurement data \( \gamma(k), u(k) \)**

   do until \( L(\theta) \) is converged to stationary value

2. **install system model matrices \( F, A, B, G, H, D, V_\nu, V_v \) and \( V_{\nu\nu} \)**

   do \( k = 0 \) until \( N - 1 \)

3. **calculate nonstationary Kalman filter described by Eqs. (3.93) to (3.100) **

4. **calculate the negative logarithm of the likelihood function Eq. (3.88) **

5. **calculate the sensitivities of the Kalman filter Eqs. (3.123) to (3.131) **

6. **calculate gradient Eq. (3.107) and information matrix Eq. (3.120) **

7. **calculate the eigenvalues and -vectors of the information matrix **

8. **is the information matrix ill conditioned?**

   - **no**
   - **yes**

9. **eliminate the smallest eigenvalues Eq. (3.164) **

10. **compute the parameter step \( \Delta \theta_i = R_i^{-1} g_i \)**

11. **one dimensional search for finding \( \alpha(i) \) in Eq. (3.104), such that \( L(\theta) \) is minimal **

12. **new estimates of parameters Eq. (3.104) **

13. **final values of estimated parameters **

14. **stop **

---

**Fig. 4.3 Program flow chart of the nonsteady state ML estimation algorithm.**
1. read initial parameter values, input and output measurement data  $y(k), u(k)$
do until $L(\theta)$ is converged to stationary value

2. instal system model matrices $F, A, B, G, H, D, V_w, V_r$ and $V_{wv}$

3. steady state Riccati equation (3.136), Kalman gain (3.137) and steady state Lyapunov equations (3.139) calculations
   do $k = 0$ until $N - 1$

4. state estimation, using constant Kalman gain

5. calculate the negative logarithm of the likelihood function Eq. (3.132)

6. calculate the sensitivities of the steady state Kalman filter

7. gradient Eq. (3.143) and information matrix Eq. (3.144) calculations

8. calculate the eigenvalues and -vectors of the information matrix

9. is the information matrix ill conditioned?
   no
   yes

10. eliminate the smallest eigenvalues Eq. (3.164)

11. compute the parameter step $\Delta \theta_i = R^{-1} _i g_i$

12. one dimensional search for fining $\alpha(1)$ Eq. (3.104), such that $L(\theta)$ is minimal

13. new estimates of parameters Eq. (3.104)

14. final values of estimated parameters

15. stop

---

Fig. 4.4 Program flow chart of the steady state ML estimation algorithm, solving the steady state Riccati and Lyapunov equations.
1. read initial parameter values, input and output measurement data \( y(k), u(k) \)

do until \( L(\theta) \) is converged to stationary value

2. install system model matrices \( F_\infty, A, B, H, D, K_P, K \)

do \( k = 0 \) until \( N-1 \)

3. state estimation using the costant matrix \( K_P \) anf \( K \)

4. calculate the negative logarithm of the likelihood function

5. calculate of the covariance matrix of the one stage prediction errors Eq. (3.147)

6. calculate the sensitivities of the Kalman filter Eqs. (3.149) (3.151), (3.153), (3.155), (3.156), (3.157), (3.159), (3.160) and (3.161)

7. calculate the gradient of the likelihood function Eqs. (3.148), (3.150) and (3.152)

8. calculate the information matrix (3.162)

9. calculate the eigenvalues and -vectors of the information matrix

10. is the information matrix ill conditioned?

\[ \text{no} \]

11. eliminate the smallest eigenvalues Eq. (3.164)

12. compute the parameter step \( \Delta \theta \) \( = \frac{1}{\nabla \theta} \) \( \theta \)

13. one dimensional search for finding \( \alpha (i) \) in Eq. (3.104) such that \( L(\theta) \) is minimum

14. new estimates of parameters, Eq.(3.104)

15. the final values of estimated parameters

16. stop

Fig. 4.5 Program flow chart of the steady state ML estimation algorithm, without solving the Riccati and Lyapunov equations.
4.3 Evaluation of Computer Programs in a Simple Simulation Test Case

To verify the developed computer programs, test calculations were performed. A second order linear system was selected to obtain ML parameter estimates, using the three algorithms described in Fig. 4.2.

In this section two test examples associated with the system without and with the correlated process and measurement noises are discussed separately.

4.3.1 Test example without correlated process and measurement noises

At first the following test system was used:

\[
\begin{align*}
\begin{bmatrix}
x_1 \\
x_2 \\
y_1 \\
y_2
\end{bmatrix} &= \begin{bmatrix}
0 & \theta_1 \\
0 & 0
\end{bmatrix} \begin{bmatrix}
x_1 \\
x_2
\end{bmatrix} + \begin{bmatrix}
u_1 \\
v_2
\end{bmatrix} + \begin{bmatrix}
u_1 \\
u_2
\end{bmatrix} w, \\
\end{align*}
\]  

The process noise \(w\) and measurement noises \(v_1\) and \(v_2\) were assumed to be white, zero-mean and to have a Gaussian distribution. The covariance matrices of these noise processes are given as follows:

\[
E \{ w(t) w^T(\tau) \} = \begin{bmatrix}
\sigma_w^2 & 0 \\
0 & \sigma_w^2
\end{bmatrix} \delta(t-\tau) \]  

\[
E \{ y(t) v^T(\tau) \} = \begin{bmatrix}
\sigma_{v_1}^2 & 0 \\
0 & \sigma_{v_2}^2
\end{bmatrix} \delta(t-\tau) \]
A pseudorandom Gaussian number generator was used to create the state and measurement noise signals. Forty seconds of data at 10 samples per second were used for this measurement generation. The first algorithm to be used for the parameter estimation of this system was the algorithm solving the nonsteady state Kalman filter. The parameters to be estimated in this example were:

\[ \theta = \begin{bmatrix} \theta_1 & \theta_2 & \sigma_w^2 & \sigma_{v_1}^2 & \sigma_{v_2}^2 \end{bmatrix}^T, \quad (4.5) \]

The input signal \( u \) was chosen as the block wave signal, as shown in Fig. 4.6. The results from this algorithm are given in Tab. 4.1.

<table>
<thead>
<tr>
<th>parameter to be estimated</th>
<th>true values</th>
<th>initial estimates</th>
<th>ML estimates</th>
<th>absolute values of errors</th>
<th>Cramer-Rao lower bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \theta_1 )</td>
<td>1.0</td>
<td>2.0</td>
<td>9.736 \times 10^{-1}</td>
<td>2.640 \times 10^{-2}</td>
<td>3.771 \times 10^{-2}</td>
</tr>
<tr>
<td>( \theta_2 )</td>
<td>1.0 \times 10^{-2}</td>
<td>2.0 \times 10^{-2}</td>
<td>1.066 \times 10^{-2}</td>
<td>6.600 \times 10^{-4}</td>
<td>5.387 \times 10^{-4}</td>
</tr>
<tr>
<td>( \sigma_w^2 )</td>
<td>1.0 \times 10^{-4}</td>
<td>2.0 \times 10^{-4}</td>
<td>9.246 \times 10^{-5}</td>
<td>7.540 \times 10^{-6}</td>
<td>2.684 \times 10^{-5}</td>
</tr>
<tr>
<td>( \sigma_{v_1}^2 )</td>
<td>1.6 \times 10^{-3}</td>
<td>4.0 \times 10^{-3}</td>
<td>1.491 \times 10^{-3}</td>
<td>1.090 \times 10^{-4}</td>
<td>1.058 \times 10^{-4}</td>
</tr>
<tr>
<td>( \sigma_{v_2}^2 )</td>
<td>1.6 \times 10^{-5}</td>
<td>4.0 \times 10^{-5}</td>
<td>1.443 \times 10^{-5}</td>
<td>1.570 \times 10^{-6}</td>
<td>1.173 \times 10^{-6}</td>
</tr>
</tbody>
</table>

Tab. 4.1 Results from the algorithm solving nonsteady state Kalman filter.

Four iterations were needed to make the negative logarithm of the likelihood function stationary and the final value of the negative logarithm of the likelihood function was:
\[ L(\theta) = -6.158 \times 10^3 \] \hspace{1cm} (4.6)

For the purpose of comparison between the three algorithms, it should be mentioned that the CPU-time on the Gould-S.E.L. 32/87 mini computer was 45.4 sec.

With the same measurement data, the second algorithm, solving steady state Riccati and Lyapunov equations, was used to estimate the same set of parameters in the same model, described by Eqs. (4.1), (4.2), (4.3) and (4.4). The results are shown in Tab. 4.2.

<table>
<thead>
<tr>
<th>parameter to be estimated</th>
<th>true values</th>
<th>initial estimates</th>
<th>ML estimates</th>
<th>absolute values of errors</th>
<th>Cramer-Rao lower bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \theta_1 )</td>
<td>1.0</td>
<td>2.0</td>
<td>9.899x10^{-1}</td>
<td>1.010x10^{-2}</td>
<td>4.361x10^{-2}</td>
</tr>
<tr>
<td>( \theta_2 )</td>
<td>1.0x10^{-2}</td>
<td>2.0x10^{-2}</td>
<td>1.062x10^{-2}</td>
<td>6.200x10^{-4}</td>
<td>5.304x10^{-4}</td>
</tr>
<tr>
<td>( \sigma_w^2 )</td>
<td>1.0x10^{-4}</td>
<td>2.0x10^{-4}</td>
<td>8.822x10^{-5}</td>
<td>1.178x10^{-5}</td>
<td>2.610x10^{-5}</td>
</tr>
<tr>
<td>( \sigma_v^2 )</td>
<td>1.6x10^{-3}</td>
<td>4.0x10^{-3}</td>
<td>1.504x10^{-3}</td>
<td>9.600x10^{-5}</td>
<td>1.068x10^{-4}</td>
</tr>
<tr>
<td>( \sigma_{v1}^2 )</td>
<td>1.6x10^{-5}</td>
<td>4.0x10^{-5}</td>
<td>1.481x10^{-5}</td>
<td>1.190x10^{-6}</td>
<td>1.198x10^{-6}</td>
</tr>
<tr>
<td>( \sigma_{v2}^2 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Tab. 4.2 Results from the algorithm solving the steady state Riccati and Lyapunov equations.

Again four iterations were used to make the negative logarithm of the likelihood function stationary and the final value of the negative logarithm of the likelihood function was:

\[ L(\theta) = -6.157 \times 10^3 \] \hspace{1cm} (4.7)
The CPU-time for running this program was 21.2 seconds on the same computer. The third algorithm is a simplified algorithm, without the solving Riccati and Lyapunov equations. In this case, the Kalman gain $K$ has to be estimated as a parameter matrix and the covariance matrices of the process and measurement noises can not be estimated directly. The parameters to be estimated for this algorithm were:

$$P = [\theta_1 \theta_2 K_{11} K_{12} K_{21} K_{22}]^T,$$  \hspace{1cm} (4.8)

where:

$$K = \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix}.$$  \hspace{1cm} (4.9)

In this algorithm, six parameters had to be estimated instead of five parameters in the previous algorithms. The results of this algorithm are given in Tab. 4.3.

Four iterations were needed to make the negative logarithm of the likelihood function stationary and the final value of the negative logarithm of the likelihood function was:

$$L(\theta) = -6.162 \times 10^3.$$  \hspace{1cm} (4.10)

The CPU-time for running this program was 15.4 seconds. Comparing this algorithm with the algorithm solving the nonsteady state Kalman filter, the third algorithm needs only less than half of the CPU-time required as the first.

The initial values of the Kalman gain matrix $K$ in Tab. 4.3 were taken from the algorithm solving the steady state Riccati and Lyapunov equations, at the initial calculation of $K$ using the initial estimates of $V_w$ and $V_v$. 
<table>
<thead>
<tr>
<th>parameter to be estimated</th>
<th>true values</th>
<th>initial estimates</th>
<th>ML estimates</th>
<th>absolute values of errors</th>
<th>Cramer-Rao lower bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta_1$</td>
<td>1.0</td>
<td>2.0</td>
<td>9.808 x 10^{-1}</td>
<td>1.920 x 10^{-2}</td>
<td>4.570 x 10^{-2}</td>
</tr>
<tr>
<td>$\theta_2$</td>
<td>1.0 x 10^{-2}</td>
<td>2.0 x 10^{-2}</td>
<td>1.064 x 10^{-2}</td>
<td>6.400 x 10^{-4}</td>
<td>5.314 x 10^{-4}</td>
</tr>
<tr>
<td>$K_{11}$</td>
<td>1.959 x 10^{-2}</td>
<td></td>
<td>8.611 x 10^{-3}</td>
<td></td>
<td>1.014 x 10^{-2}</td>
</tr>
<tr>
<td>$K_{12}$</td>
<td>1.636 x 10^{-1}</td>
<td></td>
<td>-5.803 x 10^{-3}</td>
<td></td>
<td>7.013 x 10^{-2}</td>
</tr>
<tr>
<td>$K_{21}$</td>
<td>1.636 x 10^{-3}</td>
<td></td>
<td>2.472 x 10^{-3}</td>
<td></td>
<td>3.894 x 10^{-3}</td>
</tr>
<tr>
<td>$K_{22}$</td>
<td>1.992 x 10^{-1}</td>
<td></td>
<td>2.194 x 10^{-1}</td>
<td></td>
<td>3.119 x 10^{-2}</td>
</tr>
</tbody>
</table>

Tab. 4.3 Results from the algorithm without solving the Riccati and Lyapunov equations.

From the algorithm without solving the Riccati and Lyapunov equations, the covariance matrix of prediction errors $\tilde{V}_e$ can be calculated, see Eq. (3.147), by using:

$$\tilde{V}_e = \frac{1}{N} \sum_{k=0}^{N-1} \tilde{Y}(k+1|k) \tilde{Y}^T(k+1|k).$$  \hspace{1cm} (4.11)

The final estimate of $\tilde{V}_e$ was obtained as:

$$\tilde{V}_e = \begin{bmatrix}
1.502 \times 10^{-3} & -9.347 \times 10^{-6} \\
-9.347 \times 10^{-6} & 1.846 \times 10^{-5}
\end{bmatrix}.$$  \hspace{1cm} (4.12)
As mentioned earlier, the algorithm without solving the Riccati and Lyapunov equations can not directly obtain the covariance matrices of the process and measurement noises, i.e. \( \Sigma_w \), \( \Sigma_v \), and \( \Sigma_{wv} \). However, these matrices can be reconstructed from the final results of the estimated system parameters. This can be described as follows.

From the Kalman filter algorithm, as explained in Chapter 3, using the Kalman gain matrix \( K \) and the covariance matrix of the prediction errors \( \Sigma_e \), the covariance matrix of the measurement noises can be reconstructed from Eq. (3.133) as:

\[
\Sigma_v = \Sigma_e - H K \Sigma_e
\]

Using the estimated one stage prediction gain \( K_p \), the reconstructed covariance matrix of the measurement noises \( \Sigma_{w} \), and the noise input distribution matrix \( \Sigma_w \), the covariance matrix of the correlated noises \( \Sigma_{wv} \) can be reconstructed from Eq. (3.94) as:

\[
\Sigma_{wv} = (\Sigma_w \Sigma_w)^{-1} \Sigma_w K_p \Sigma_v
\]

Using the results shown in Tab. 4.3, the diagonal elements of the covariance matrix of the measurement noise were as:

\[
\sigma^2_{v_1} = 1.489 \times 10^{-3},
\]

\[
\sigma^2_{v_2} = 1.443 \times 10^{-5},
\]

and \( \Sigma_{wv} = 0 \), since \( K_p \) was assumed to be zero, which means that the process and the measurement noises were assumed to be independent.

The results, as shown in Eqs. (4.15) and (4.16), can be compared to the results described in Tab. 4.2, in which the covariance matrix of the measurement noises was obtained from the algorithm solving the steady state Riccati and Lyapunov equations.
The re-construction of the covariance matrix of the process noises in this algorithm can be performed as follows.

Rewriting the one stage prediction algorithm and the measurement update algorithm as described by Eqs. (3.92) and (3.95), results in:

$$\hat{x}(k+1|k) = \Phi \hat{x}(k|k) + \Gamma_u u(k) + K_p [y(k) - H \hat{x}(k|k) - D u(k)] \quad (4.17)$$

$$\hat{x}(k+1|k+1) = \hat{x}(k|k) + K [y(k+1) - H \hat{x}(k+1|k) - D u(k+1)] \quad (4.18)$$

From substituting (4.18) into (4.17), it follows that:

$$\hat{x}(k+1|k+1) = \Phi \hat{x}(k|k) + \Gamma_u u(k) + \left( - K_p H - K H \Phi + K H K_p H \right) \hat{x}(k|k) +$$

$$+ \left( - K_p D - K H \Gamma_u + K H K_p D \right) u(k) - K D u(k+1) +$$

$$+ \left[ K_p - K H K_p \right] y(k) + K y(k+1) \quad (4.19)$$

Comparing (4.19) to the original system dynamic model:

$$x(k+1) = \Phi x(k) + \Gamma_u u(k) + \Gamma_w w(k) \quad (4.20)$$

it can be found that the process noise vector follows from the extra terms of (4.19) and then the process noise vector $\hat{w}$ can be re-constructed as:

$$\hat{w}(k) = \left( \Gamma_w^T \Gamma_w \right)^{-1} \Gamma_w^T \left[ - K_p H - K H \Phi + K H K_p H \right] \hat{x}(k|k) +$$

$$+ \left( - K_p D - K H \Gamma_u + K H K_p D \right) u(k) - K D u(k+1) +$$

$$+ \left[ K_p - K H K_p \right] y(k) + K y(k+1) \quad (4.21)$$
Therefore, the covariance matrix of the estimated process noises can be reconstructed as:

\[
\hat{V}_w = \frac{1}{N} \sum_{k=1}^{N} \hat{w}(k) \hat{w}^T(k). \tag{4.22}
\]

Using the results obtained in this run, the reconstructed covariance matrix of the process noises turned out as:

\[
\hat{V}_w = \hat{\sigma}_w^2 = 8.828 \times 10^{-5}. \tag{4.23}
\]

It should be pointed out, that from the algorithm without solving the Riccati and Lyapunov equations, the Cramer-Rao low bound of the covariance matrices \(V_w, V_v, V_{wv}\) cannot be obtained. However, the Cramer-Rao low bound of the estimated one stage prediction gain matrix \(K_p\) and the Kalman gain matrix \(K\) can be obtained, see Tab. 4.3.

The time response of the state estimates of the three algorithms were nearly the same. They are shown in Fig. 4.7 to 4.14.

![Graph](image)

Fig. 4.6 Input signal \(u(k)\).
**Fig. 4.7** Time responses of the true $x_1(k)$ and the estimated $\hat{x}_1(k|k)$.

**Fig. 4.8** Time responses of the true $x_2(k)$ and the estimated $\hat{x}_2(k|k)$.

**Fig. 4.9** Measurement data $y_1(k)$.

**Fig. 4.10** Measurement data $y_2(k)$. 
Fig. 4.11 Prediction error $\tilde{y}_1(k+1|k)$.

Fig. 4.12 Prediction error $\tilde{y}_2(k+1|k)$.

Fig. 4.13 Estimated process noise $\hat{w}(k)$.

Fig. 4.14 Simulated process noise $w(k)$. 
4.3.2 Test example with correlated process and measurement noises

The second test case used the same dynamical model, Eq. (4.1) and one more measurement, \( y_3 = x_2 \) was used for parameter estimation. The measurement equation (4.2) then became:

\[
\begin{align*}
\dot{x} &= C_1 x + C_2 \dot{x} + v \\
&= H x + D u + \nu^* 
\end{align*}
\]

(4.24)

where:

\[
C_1 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix},
\]

\[
C_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix},
\]

\[
H = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix},
\]

\[
D = [0 \ 0 \ \theta_2]^T,
\]

and:

\[
\nu^* = [v_1 \ v_2 \ (w + v_3)]^T,
\]

therefore:
\[
\begin{bmatrix}
y_1 \\
y_2 \\
y_3
\end{bmatrix} =
\begin{bmatrix}
1 & 0 \\
0 & 1 \\
0 & 0
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
x_3
\end{bmatrix} +
\begin{bmatrix}
0 \\
0 \\
\theta_2
\end{bmatrix}
\begin{bmatrix}
u_1 \\
u_2 \\
u_3
\end{bmatrix} +
\begin{bmatrix}
0 \\
w \\
0
\end{bmatrix} +
\begin{bmatrix}
0 \\
0 \\
1
\end{bmatrix}
\begin{bmatrix}
v_1 \\
v_2 \\
v_3
\end{bmatrix}.
\]

The additional measurement \( y_3 \) can be viewed as an acceleration measurement of the system. The parameters to be estimated were extended to:

\[
\hat{\theta} = [\theta_1 \quad \theta_2 \quad \sigma_w^2 \quad \sigma_{v_1}^2 \quad \sigma_{v_2}^2 \quad \sigma_{v_3}^2]^T, \quad (4.25)
\]

where:

\[
E \{ y(t) y^T(t) \} = \begin{bmatrix}
\sigma_{v_1}^2 & 0 & 0 \\
0 & \sigma_{v_2}^2 & 0 \\
0 & 0 & \sigma_{v_3}^2
\end{bmatrix} \delta(t-t_0), \quad (4.26)
\]

Since measurement \( y_3 \) contained the process noise \( w \), the measurement noise of \( y_3 \) was then correlated with the process noise. Therefore, the estimation algorithm with correlated process and measurement noise had to be used for this system. The purpose of this example was to test the estimation algorithm with correlated process and measurement noise. This is of interest in view of the acceleration measurement being part of the measurements on the flexible substructures of the spacecraft. Only the algorithm solving the steady state Riccati and Lyapunov equations was applied in this test. The results are given in Tab. 4.4.

To re-construct the response of \( y_3 \), the process noise \( w \) again needed to be re-constructed using Eq. (4.21), since from Eq. (4.24) it follows that:

\[
\hat{y}_3 = \hat{\theta}_2 u + \hat{w}. \quad (4.27)
\]

The state estimation from the above results are shown in Figs. 4.15 to 4.19.
<table>
<thead>
<tr>
<th>parameter to be estimated</th>
<th>true values</th>
<th>initial estimates</th>
<th>ML estimates</th>
<th>absolute values of errors</th>
<th>Cramer-Rao lower bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta_1$</td>
<td>1.0</td>
<td>2.0</td>
<td>9.778x10^-1</td>
<td>2.220x10^-2</td>
<td>4.321x10^-2</td>
</tr>
<tr>
<td>$\theta_2$</td>
<td>1.0x10^-2</td>
<td>2.0x10^-2</td>
<td>1.048x10^-2</td>
<td>4.800x10^-4</td>
<td>5.011x10^-4</td>
</tr>
<tr>
<td>$\sigma_w^2$</td>
<td>1.0x10^-4</td>
<td>2.0x10^-4</td>
<td>9.610x10^-5</td>
<td>3.900x10^-6</td>
<td>8.263x10^-6</td>
</tr>
<tr>
<td>$\sigma_{v_1}^2$</td>
<td>1.6x10^-3</td>
<td>4.0x10^-3</td>
<td>1.498x10^-3</td>
<td>1.020x10^-4</td>
<td>1.063x10^-4</td>
</tr>
<tr>
<td>$\sigma_{v_2}^2$</td>
<td>1.6x10^-5</td>
<td>4.0x10^-5</td>
<td>1.493x10^-5</td>
<td>1.070x10^-6</td>
<td>1.088x10^-6</td>
</tr>
<tr>
<td>$\sigma_{v_3}^2$</td>
<td>1.6x10^-5</td>
<td>4.0x10^-5</td>
<td>6.484x10^-6</td>
<td>9.516x10^-6</td>
<td>3.936x10^-6</td>
</tr>
</tbody>
</table>

Tab. 4.4 Results from the algorithm solving the steady state Riccati and Lyapunov equations (correlated process and measurement noises due to $y_3$ involved).

Four iterations were needed to make the negative logarithm of the likelihood function stationary and the final value of the negative logarithm of the likelihood function was:

$$L(\theta) = -9.498 \times 10^3$$  \hspace{1cm} (4.28)
Fig. 4.15 Time responses of the true $x_1(k)$ and the estimated $\hat{x}_1(k|k)$.

Fig. 4.16 Time response of the true $x_2(k)$ and the estimated $\hat{x}_2(k|k)$.

Fig. 4.17 the True $\theta_2 u(k) + w(k)$ and the estimated $\hat{\theta}_2 u(k) + \hat{w}(k)$.

Fig. 4.18 Measurement $y_3(k)$. 
As a summary of this section, it can be concluded that:

1. the algorithm solving the nonsteady state Riccati and Lyapunov equations, i.e. the nonsteady state Kalman filter is the slowest algorithm,
2. the algorithm solving the steady state Riccati and Lyapunov equations, i.e. the steady state Kalman filter, is faster than the algorithm solving the nonsteady state Riccati and Lyapunov equations, and
3. the algorithm without the solving Riccati and Lyapunov equations is the fastest algorithm of these three algorithms, but the number of parameters to be estimated may possibly be larger than the algorithms solving the Riccati and Lyapunov equations.

The accuracy of the estimates was satisfactory, since all the estimated parameters were closed to the true parameters and all the absolute values of the errors between the true parameters and estimated parameters were in the same order as the Cramer-Rao lower bounds, as shown in Tabs. 4.1 to 4.4. The accuracy of the predicted outputs of the system from these algorithms can be judged by the so-called goodness of fit, which defined as follows.

The goodness of fit of the $i^{th}$ predicted output can be obtained using:
\[ \text{GOF}_i = \left\{ 1 - \left[ \sum_{k=1}^{N} \frac{\tilde{y}^2(k+1|k)}{\sum_{k=1}^{N} y^2(k+1)} \right]^{1/2} \right\}, \quad (a) \]

where \( i = 1, 2, \ldots, q. \)

The overall goodness of fit can be calculated using:

\[ \text{GOF} = \left\{ 1 - \left[ \sum_{k=1}^{N} \frac{y^T(k+1|k) \tilde{y}(k+1|k)}{\sum_{k=1}^{N} y^T(k+1) y(k+1)} \right]^{1/2} \right\}. \quad (b) \]

The goodness of fit obtained from previous tests is given in Tab. 4.5.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>1</th>
<th>Goodness of fit of predicted output i (%)</th>
<th>Overall goodness of fit (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solving non-stationary Kalman filter</td>
<td>2</td>
<td>95.66</td>
<td>92.66</td>
</tr>
<tr>
<td>Solving stationary Kalman filter</td>
<td>2</td>
<td>95.66</td>
<td>92.66</td>
</tr>
<tr>
<td>Without solving Ricc. and Lyap. equations</td>
<td>2</td>
<td>95.67</td>
<td>92.70</td>
</tr>
<tr>
<td>Stationary Kalman filter with additional meas. ( y_3 )</td>
<td>3</td>
<td>95.08</td>
<td></td>
</tr>
</tbody>
</table>

Tab. 4.5 Goodness of fit of the predicted outputs to the measurements.

From Tab. 4.5, the goodness of fit of the predicted outputs are all below 96%. It is due to the fact that the goodness of fit, as calculated by Eqs. (a) and (b), depends on the covariances of the predicted errors of the outputs.
4.4 Simulation of the Flexible Spacecraft Motions and Generation of the Measurement Data

The time response of a simulated flexible spacecraft is needed to provide the measurement data. This time response should be as accurate as possible as compared to a real spacecraft and, therefore, should be generated by a high order dynamical model. The dynamical model of the spacecraft to be used in this section has been discussed in Section 2.5. The configuration of the spacecraft is shown in Fig. 4.20. To simulate the time response for generating the measurement data, the parameters of the spacecraft have to be specified. The parameters as selected are listed in Tabs. 4.6a and 4.6b. The configuration of the hypothetical satellite resembles that of some communication satellites, such as the Canadian Communication Technology Satellite (CTS) which also consists of a rigid central body and two symmetrical flexible solar panels, see Ref. 2.28. However, the model considered in the simulation of this chapter is larger than the CTS.

<table>
<thead>
<tr>
<th>parameters as selected in the simulation model</th>
<th>values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Degrees of freedom</td>
<td>3 (rotations)</td>
</tr>
<tr>
<td>Moment of inertia $I_x$</td>
<td>4480 kg m$^2$</td>
</tr>
<tr>
<td>Moment of inertia $I_y$</td>
<td>4480 kg m$^2$</td>
</tr>
<tr>
<td>Moment of inertia $I_z$</td>
<td>4480 kg m$^2$</td>
</tr>
<tr>
<td>Product of inertia $I_{xy}$</td>
<td>0 kg m$^2$</td>
</tr>
<tr>
<td>Product of inertia $I_{xz}$</td>
<td>0 kg m$^2$</td>
</tr>
<tr>
<td>Product of inertia $I_{yz}$</td>
<td>0 kg m$^2$</td>
</tr>
<tr>
<td>Dimension of rigid control body</td>
<td>$2 \times 2 \times 2$ m$^3$</td>
</tr>
</tbody>
</table>

Tab. 4.6a Parameters to be used in the simulation model (rigid body).
<table>
<thead>
<tr>
<th>Parameters as selected in the simulation model</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dimension of each solar panel</td>
<td>48 x 4 x 0.04 m³</td>
</tr>
<tr>
<td>Young's modulus E</td>
<td>5.0 x 10¹⁰ N/m²</td>
</tr>
<tr>
<td>Poisson coefficient ν</td>
<td>0.3</td>
</tr>
<tr>
<td>Mass density ρ</td>
<td>1.0 x 10³ kg/m³</td>
</tr>
<tr>
<td>Number of elements in each panel</td>
<td>16 (rectangular)</td>
</tr>
<tr>
<td>Dimension of each element (b x a x t)</td>
<td>6 x 2 x 0.04 m³</td>
</tr>
<tr>
<td>Number of degrees of freedom of each node</td>
<td>3</td>
</tr>
<tr>
<td>(one translation and two rotations)</td>
<td></td>
</tr>
<tr>
<td>Number of nodes in each element</td>
<td>4</td>
</tr>
<tr>
<td>Number of degrees of freedom in each element</td>
<td>12</td>
</tr>
<tr>
<td>Number of nodes in each panel</td>
<td>27</td>
</tr>
<tr>
<td>Number of solar panels</td>
<td>2</td>
</tr>
<tr>
<td>Offset angle of solar panel δ</td>
<td>0</td>
</tr>
<tr>
<td>Damping coefficients α and β</td>
<td>0, 0</td>
</tr>
<tr>
<td>Total degrees of freedom (2 x 3 x 27)</td>
<td>162</td>
</tr>
</tbody>
</table>

Tab. 4.6b Parameters to be used in the simulation model (flexible solae panels).

It is assumed, that the selected number of elements in the flexible solar panels is sufficiently high to obtain a reasonably accurate dynamical model of the discussed flexible spacecraft.

Using the parameters in Tab. 4.6, the symmetric stiffness matrix of (B.1) in Appendix B for the continuous shape function matrix (2.108) in Chapter 2 can be calculated as follows:
\[ K_{m1} = 2.442 \times 10^{-3} \times \]

\[
\begin{array}{cccc}
43.5 \\
39.0 & 54.5 \\
-41.5 & -38.9 & 54.8 & \text{Symmetric} \\
10.5 & 17.1 & -12.5 & 43.5 \\
-17.1 & -27.7 & 19.0 & -39.0 & 54.5 \\
-12.5 & -19.0 & 17.2 & -41.5 & 38.9 & 54.8 \\
-11.2 & -19.1 & 13.5 & -42.8 & 37.0 & 40.5 & 43.5 \\
19.1 & 31.7 & 20.1 & -46.5 & -34.3 & -39.0 & 54.5 \\
-13.5 & -20.1 & 9.62 & -40.5 & 34.3 & 26.4 & 41.5 & -38.9 & 54.8 \\
-42.8 & -37.0 & 40.5 & -11.2 & 19.1 & 13.5 & 10.5 & -17.1 & 12.5 & 43.5 \\
-37.0 & -46.5 & 34.3 & -19.1 & 31.7 & 20.1 & 17.1 & -27.7 & 19.0 & 39.0 & 54.5 \\
-40.5 & -34.3 & 26.4 & -13.5 & 20.1 & 9.62 & 12.5 & -19.0 & 17.2 & 41.5 & 38.9 & 54.8 \\
\end{array}
\]

(4.29)

and the symmetric mass matrix of (8.2) in Appendix B can be calculated for
the continuous shape function matrix (2.108) as:

\[ M_i = 2.7211 \times 10^{-3} \times \]

\[
\begin{array}{cccc}
24,336 \\
20,592 & 22,464 \\
-6,864 & -5,808 & 2,496 \\
8,424 & 12,186 & -2,376 & 24,336 \\
-12,168 & -16,848 & 3,432 & -20,592 & 22,464 \\
-2,376 & -3,432 & 864 & -6,864 & 5,808 & 2,496 \\
2,916 & 4,212 & -1,404 & 8,424 & -7,128 & -4,056 & 24,336 \\
-4,212 & -5,832 & 2,028 & -7,128 & 7,776 & 3,432 & -20,592 & 22,464 \\
1,404 & 2,028 & -640 & 4,056 & -3,432 & -1,872 & 6,864 & -5,808 & 2,496 \\
8,424 & 7,128 & -4,056 & 2,916 & -4,212 & -1,404 & 8,424 & -12,186 & 2,376 & 24,336 \\
7,128 & 7,776 & -3,432 & 4,212 & -5,832 & -2,028 & 12,168 & -16,848 & 3,432 & 20,592 & 22,464 \\
4,056 & 3,432 & -1,872 & 1,404 & -2,028 & -640 & 2,376 & -3,432 & 864 & 6,864 & 5,808 & 2,496 \\
\end{array}
\]

(4.30)
The elements of coupling matrix $A_1$ of (2.126) in Chapter 2 depend on each particular element i on the panel. For example, picking up the first element of the right hand side solar panel shown in Fig. 4.20 and using Eqs. (2.114) and (2.115) in Chapter 2, it follows that:

$$y_{01} = 1 = 1 \text{ m ,}$$  \hspace{1cm} (4.31)

and:

$$x_{01} = -a = -2 \text{ m .}$$  \hspace{1cm} (4.32)

The coupling matrix $A_1$ can then be calculated from Eq. (2.119) as:

$$A_1 = \begin{bmatrix} 16.8 & 8.4 & -9.2 & 31.2 & -27.6 & -10.4 & 31.2 & -27.6 & 10.4 & 16.8 & 20.4 & 5.6 \\ 10.8 & 8.4 & -2.8 & 8.4 & -8.4 & -2.4 & 3.6 & -3.6 & 1.6 & 3.6 & 3.6 & 1.6 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$A_1 = \begin{bmatrix} 10.8 & 8.4 & -2.8 & 8.4 & -8.4 & -2.4 & 3.6 & -3.6 & 1.6 & 3.6 & 3.6 & 1.6 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$  \hspace{1cm} (4.33)

The inertia matrix of the first element of the right hand solar panel can be written according to (2.122) as:

$$I_{11} = \begin{bmatrix} 19 & 4 & 0 \\ 4 & 1.333 & 0 \\ 0 & 0 & 20.333 \end{bmatrix}$$  \hspace{1cm} (4.34)

A computer program for calculating the inertia and coupling matrices of all elements and assembling all the matrices in the total dynamical model of the spacecraft is needed. Since the rotation offset angle of the solar panels is assumed to be zero, the transformation matrix $T_{r}$ in (2.92) will be a $3 \times 3$ unit matrix, i.e., $T_{r} = U_3$ for the right hand side solar panel. For the left hand side solar panel, the transformation matrix $T_{l}$ in (2.93) will be the form:
\[
\mathbf{I}_1 = \begin{bmatrix}
-1 & 0 & 0 \\
0 & -1 & 0 \\
0 & 0 & 1
\end{bmatrix}.
\]

Therefore, the total inertia matrix of the flexible solar panels can be simply written as:

\[
\mathbf{I}_{\text{a}} = \frac{N}{2} \sum_{i=1}^{N/2} \mathbf{I}_i + \sum_{i=N/2+1}^{N} \mathbf{T}_{i}^T \mathbf{I}_i \mathbf{T}_{i}.
\]  

(4.35)

Whereas the total coupling matrix \( \mathbf{A} \) in (2.62) can be written as:

\[
\mathbf{A} = \frac{N}{2} \sum_{i=1}^{N/2} \mathbf{A}_i \mathbf{P}_i + \sum_{i=N/2+1}^{N} \mathbf{T}_{i}^T \mathbf{A}_i \mathbf{P}_i.
\]  

(4.36)

The assembling matrix \( \mathbf{P}_i \), as defined in Eq. (2.56), depends on the choice of the discussed element. Since each element is selected as four nodes on the corners of the rectangular plate element, see Chapter 2, the displacement vector of element \( i \) can be written as:

\[
d_i = [d_{i,1}^T, d_{i,2}^T, d_{i,3}^T, d_{i,4}^T]^T.
\]

Each node has three degrees of freedom (one translation and two rotations), see also Tab. 4.1, i.e.:

\[
d_{i,k} = [d_{i,k,1}, d_{i,k,2}, d_{i,k,3}]^T.
\]

where \( d_{i,k} \) denotes the displacement vector at the \( k^{th} \) node of the \( i^{th} \) element, \( d_{i,k,1} \) is the translational displacement of the node along the z axis of the local reference frame, \( d_{i,k,2} \) and \( d_{i,k,3} \) are the rotational displacements of the node about the x and y axes of the local reference frame.
In the simulation the total number of nodes in the solar panels combined is 54 and the total displacement vector \( d \) is then a 162 x 1 vector \( (162 = 3 \times 54) \), arranged as follows:

\[
d = [(d_1^*)^T \ (d_2^*)^T \ \ldots \ (d_{k-1}^*)^T \ (d_k^*)^T \ (d_{k+1}^*)^T \ \ldots \ (d_{54}^*)^T]^T,
\]

in which \( d_k^* \) is the displacement vector of the \( k \)th node arranged in the total displacement vector \( d \).

The relation between the element displacement vector \( d_1 \) and the total displacement vector \( d \) can be found from Eq. (2.56) in Chapter 2, that is:

\[
d_1 = P_{11} \ d.
\]

The assembling matrix depends on the choice of the element. For example, the assembling matrix \( P_{11} \) of the first element of the right hand solar panel, see Fig. 4.20, can then be written as:

\[
P_{11} = \begin{bmatrix}
0_3 & 0_3 & 0_3 & 0_3 & 0_3 & 0_3 & \ldots & 0_3 & 0_3 & 0_3 & \ldots & 0_3 \\
0_3 & 0_3 & 0_3 & 0_3 & 0_3 & 0_3 & \ldots & 0_3 & 0_3 & 0_3 & \ldots & 0_3 \\
0_3 & 0_3 & 0_3 & 0_3 & 0_3 & 0_3 & \ldots & 0_3 & 0_3 & 0_3 & \ldots & 0_3 \\
0_3 & 0_3 & 0_3 & 0_3 & 0_3 & 0_3 & \ldots & 0_3 & 0_3 & 0_3 & \ldots & 0_3 \\
0_3 & 0_3 & 0_3 & 0_3 & 0_3 & 0_3 & \ldots & 0_3 & 0_3 & 0_3 & \ldots & 0_3 \\
1 & 2 & 3 & 4 & 5 & 6 & \ldots & k-1 & k & k+1 & \ldots & 54
\end{bmatrix}
\]

(sequence number of node)

Moreover, the assembling matrix \( P_{11} \) has the following property, see Eqs. (2.57) and (4.37):

\[
P_{11} \cdot P_{11}^T = U_{12}.
\]
Next, input signals to be used in the parameter estimation experiments should be selected. The importance of choosing appropriate test inputs for exciting a system has been recognized for a long time, see e.g. [3.45] and [3.46]. Several considerations related to the selection of inputs are:

1. The input should be capable of being implemented easily through the actuators on board the spacecraft and they should be easily measured.
2. The resulting response of the spacecraft should not surpass the constraints of linearity.
3. The measured response of the spacecraft should be sensitive to the parameters to be estimated, in order that the inputs may yield good estimates of parameters.

The theory of input design for linear system identification is often formulated as an optimal control problem (see e.g. [3.45] and [3.46]). The performance criterion used is the sensitivity of the system response to the unknown parameters or some new criterion (see the work of Mulder [5.1]) defining the input signal. This subject will not be included in this report. For the flexible spacecraft parameter identification problem, only simple pulse input signals will be considered, capable of exciting the most important modes of the discussed flexible spacecraft. The lowest modes are of particular importance, since the lowest frequencies are assumed to be of special interest for control algorithm synthesis (see Chapter 3). Based on a Fourier expansion, a pulse signal may be considered as containing an infinite number of frequencies. The lowest frequency of the pulse depends on the duration of the pulse. When the lowest frequencies of the spacecraft have to be excited, the pulse signals to be used should be of sufficient duration. A pulse can be easily implemented, e.g. using gas jets which are capable of giving constant amplitude moments (or forces). Pulse shaped moments may also be obtained from reaction wheels on the spacecraft, which means that such that such actuators may also be considered as input signal generators.

In this study, two types of inputs are considered to excite the flexible spacecraft, viz. the doublet input, consisting of two pulses, and a single pulse input, see Figs. 4.21 to 4.24.

The simulated motions of the flexible spacecraft were calculated, using the Gould S.E.L MPX-32/87 computer in the Subject Group of Stability and Control of the Faculty of Aerospace Engineering, TU-Delft.
Since the offset angle of the solar panels is assumed to be zero, it can be shown that, from the resulting symmetry, the motions about the yaw axis (the z axis of the body fixed reference frame) are independent from the motions about the roll and pitch axes. It turns out, that the yaw motion is only the rigid body motion if the in-plane flexibility of the solar panels is negligible, as discussed in Chapter 2. Therefore, for the motions of flexible solar panels only the coupled motions about the roll and pitch axes (the x and y axes of the body fixed reference frame respectively) are considered.

The time responses to these inputs are shown in Fig. 4.25 to Fig. 4.40.

The selection of the measured variables is now discussed.

At first, the measured variables of the rigid control body may be considered as the attitude velocities and the attitude angles of the rigid control body. The attitude velocities may be measured by rate gyros and the attitude angle may be measured by some attitude sensors (sun sensors, star trackers, horizon sensors, magnetometers, gravity gradient sensors, etc.), as discussed in Subsection 3.2.1. Therefore, the measured variables of the rigid control body are considered as the roll and pitch velocities, and roll and pitch angles.

For the measurements of the flexible solar panels, only the accelerometers are selected. The reason of applying the accelerometers to the flexible solar panels has been discussed in Subsection 3.2.1. The accelerometers, which are placed on the flexible solar panels, in fact, measure the specific forces at those places on the flexible solar panels. The difference between the measured specific force and the required acceleration is the component of gravity along the sensitive axis of the transducer. Assuming the attitude of the transducer relative to the local vertical to vary only very slowly, fluctuations in the specific forces will be very rarely identical to the fluctuations in the kinematic accelerations. Therefore, it is assumed that the accelerometers measure kinematic acceleration rather than specific force, as discussed in Chapter 1.

In the system dynamic model, as described by Eq. (2.91) in Chapter 2, the accelerations \( \ddot{q} \) are related to the flexible substructure local reference frame, which is not the inertial reference frame. The difference between the acceleration of a node with respect to the local reference frame and the acceleration of the node with respect to the inertial reference frame is that the acceleration of the node with respect to the inertial reference
frame includes the acceleration of the undeformed spacecraft related to the inertial reference frame, of which the acceleration of the node with respect to the local reference frame does not consist. Assuming that using some additional accelerometers in the rigid main body of the spacecraft, the acceleration of the node with respect to the local reference may be measured, in the followint discussions, it is assumed that $a$ may be measured.

In the simulation experiments, the measured accelerations were selected on four tip nodes of the flexible solar panels, i.e. nodes 25, 27, 52 and 54, see Fig. 4.20. The reason is that the accelerations on the tip nodes of the solar panels are usually larger than the accelerations on the nodes closed to the rigid control body because the solar panels are fixed on the control body, see also Fig. 4.20. Two accelerometers, which were located on the tip nodes of each solar panel, provided the bending and twist motions of the flexible panel.

The numerical integration algorithm was used to calculate the time responses of the simulated flexible spacecraft dynamic model. The obtained time responses are given in Figs. 4.25 to 4.40.

![Fig. 4.20 Configuration of the simulated spacecraft.](image)

To generate the measurement data for ML parameter estimation, the previous time responses were corrupted by some measurement noises, as discussed in the introduction of this chapter. These noises can be assumed as a simulation of the instrumental measurement errors. A pseudorandom number generator was used to create the measurement noise signals. The standard deviations of
these simulated noises were chosen, according to the amplitudes of the simulated time responses of the motions of the satellite in Figs. 4.25 to 4.40. The covariance matrix \( \Phi \) was chosen as a diagonal matrix, which means that the measurement noises were assumed to be independent. The selected values of the diagonal elements of the covariance matrix of the measurement noises \( \Phi \) are given in Tab. 4.7.

Fig. 4.21 Roll axis input (doublet).

Fig. 4.22 Pitch axis input (doublet).

Fig. 4.23 Roll axis input (single pulse).

Fig. 4.24 Pitch axis input (single pulse).
Fig. 4.25 Simulated roll angular rate (doublet).

Fig. 4.26 Simulated roll angle (doublet).

Fig. 4.27 Simulated roll angular rate (single pulse).

Fig. 4.28 Simulated roll angle (single pulse).
Fig. 4.29 Simulated pitch angular rate (doublet).

Fig. 4.30 Simulated pitch angle (doublet).

Fig. 4.31 Simulated pitch angular rate (single pulse).

Fig. 4.32 Simulated pitch angle (single pulse).
Fig. 4.33 Simulated acceleration of node 25 (doublet).

Fig. 4.34 Simulated acceleration of node 27 (double).

Fig. 4.35 Simulated acceleration of node 25 (single pulse).

Fig. 4.36 Simulated acceleration of node 27 (single pulse).
Fig. 4.37 Simulated acceleration of node 52 (doublet).

Fig. 4.38 Simulated acceleration of node 54 (doublet).

Fig. 4.39 Simulated acceleration of node 52 (single pulse).

Fig. 4.40 Simulated acceleration of node 54 (single pulse).
Square roots of diagonal elements of $V_V$

<table>
<thead>
<tr>
<th>Element Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma_{V11}$ (acceleration of node 25)</td>
<td>$3.8 \times 10^{-3}$ m/s$^2$</td>
</tr>
<tr>
<td>$\sigma_{V22}$ (acceleration of node 27)</td>
<td>$3.8 \times 10^{-3}$ m/s$^2$</td>
</tr>
<tr>
<td>$\sigma_{V33}$ (acceleration of node 52)</td>
<td>$3.8 \times 10^{-3}$ m/s$^2$</td>
</tr>
<tr>
<td>$\sigma_{V44}$ (acceleration of node 54)</td>
<td>$3.8 \times 10^{-3}$ m/s$^2$</td>
</tr>
<tr>
<td>$\sigma_{V55}$ (roll velocity of control body)</td>
<td>$4.8 \times 10^{-3}$ rad/s</td>
</tr>
<tr>
<td>$\sigma_{V66}$ (pitch velocity of control body)</td>
<td>$4.8 \times 10^{-3}$ rad/s</td>
</tr>
<tr>
<td>$\sigma_{V77}$ (roll angle of control body)</td>
<td>$4.8 \times 10^{-3}$ rad</td>
</tr>
<tr>
<td>$\sigma_{V88}$ (pitch angle of control body)</td>
<td>$4.8 \times 10^{-3}$ rad</td>
</tr>
</tbody>
</table>

Tab. 4.7. The selected values of the diagonal elements of the covariance matrix of the measurement noises as used for simulation.

4.5 A Priori Parameter Value Selections

To start the maximum likelihood parameter estimation program, initial parameter values are required. The steps for choosing the initial parameter values may be as follows.

1. The initial parameter values resulting from the dimensions of the elements can be readily obtained from the design drawings of the spacecraft. When the material of the flexible appendages is known from the design of the spacecraft, the mass density $\rho$, Young's modulus $E$ and Poisson ratio $\nu$ can be chosen from ground experiments.
2. The choice of the initial parameters relating to the measurement instruments (attitude, attitude velocity and acceleration measurement devices) and the initial parameter values of the covariance matrix of the measurement noises can be determined from ground calibration tests of the measurement instruments.

3. The stochastic model for the modelling errors should be discussed next. The model of the modelling errors first may be assumed to consist of white noise processes. The modelling error model discussed in Subsection 3.2.3 can be used when the results of the parameter and state estimates are not satisfactory. For simplicity, the matrix \( E \) defined in Eq. (3.43) can be chosen as a diagonal matrix with negative values on all its diagonal elements, because of the required stability.

4. Initial estimates of required input signal levels can be determined from ground tests of the actuators to be used on the spacecraft and from a subsequent numerical simulation of the spacecraft's response.

Using the foregoing considerations, initial estimates of the structural parameters are given in Tab. 4.8. Parameters are deliberately offset from the true values, as foll in Tab. 4.6. The dimensions of the spacecraft are also given in the following table.

The modeling error estimation entails the choice between the two modelling error models and their unknown parameters. If the modeling error is chosen as a white noise process, with zero mean and having a Gaussian distribution, the unknown parameters are just the elements of the covariance matrix of this white noise, i.e. of \( w_w \), and the order of the system model will not be increased. However, the order of the total system, including the error model, will be increased if the modelling errors are considered as nonwhite process noises and the unknown parameters exist in both the covariance matrix \( w_w \) and the matrix \( E \), see Eq. (3.43). In Section 4.6 the two types of modelling errors will be discussed for the reduced spacecraft model and the initial parameters will be discussed for each modelling error model, viz. one model with white noise and one with coloured noise.
<table>
<thead>
<tr>
<th>Parameters to be estimated</th>
<th>Initial estimates of parameters</th>
<th>True values from Tab. 4.6</th>
</tr>
</thead>
<tbody>
<tr>
<td>$I_x$</td>
<td>$4400 \text{ kg m}^2$</td>
<td>$4480 \text{ kg m}^2$</td>
</tr>
<tr>
<td>$I_y$</td>
<td>$4400 \text{ kg m}^2$</td>
<td>$4480 \text{ kg m}^2$</td>
</tr>
<tr>
<td>$\rho$</td>
<td>$1.0 \times 10^3 \text{ kg m}^3$</td>
<td>$1.0 \times 10^3 \text{ kg m}^3$</td>
</tr>
<tr>
<td>$E$</td>
<td>$6.0 \times 10^{10} \text{ N/m}^2$</td>
<td>$5.0 \times 10^{10} \text{ N/m}^2$</td>
</tr>
<tr>
<td>$v$</td>
<td>0.4</td>
<td>0.3</td>
</tr>
<tr>
<td>$a$</td>
<td>6 m</td>
<td>6 m</td>
</tr>
<tr>
<td>$b$</td>
<td>2 m</td>
<td>2 m</td>
</tr>
<tr>
<td>$t$</td>
<td>0.04 m</td>
<td>0.04 m</td>
</tr>
<tr>
<td>$h$</td>
<td>220 Nm</td>
<td>200 N m</td>
</tr>
</tbody>
</table>

Tab. 4.8 Initial parameter values to be used for starting the ML parameter estimation program.
The initial values of the covariance matrix of the measurement noises are given in Tab. 4.9.

<table>
<thead>
<tr>
<th>Square roots of diagonal elements of $V_v$</th>
<th>Initial estimates</th>
<th>True values given in Tab. 4.7</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma_{v11}$</td>
<td>$3.0 \times 10^{-3}$ m/s$^2$</td>
<td>$3.8 \times 10^{-3}$ m/s$^2$</td>
</tr>
<tr>
<td>$\sigma_{v22}$</td>
<td>$3.0 \times 10^{-3}$ m/s$^2$</td>
<td>$3.8 \times 10^{-3}$ m/s$^2$</td>
</tr>
<tr>
<td>$\sigma_{v33}$</td>
<td>$3.0 \times 10^{-3}$ m/s$^2$</td>
<td>$3.8 \times 10^{-3}$ m/s$^2$</td>
</tr>
<tr>
<td>$\sigma_{v44}$</td>
<td>$3.0 \times 10^{-3}$ m/s$^2$</td>
<td>$3.8 \times 10^{-3}$ m/s$^2$</td>
</tr>
<tr>
<td>$\sigma_{v55}$</td>
<td>$4.0 \times 10^{-3}$ rad/s</td>
<td>$4.8 \times 10^{-3}$ rad/s</td>
</tr>
<tr>
<td>$\sigma_{v66}$</td>
<td>$4.0 \times 10^{-3}$ rad/s</td>
<td>$4.8 \times 10^{-3}$ rad/s</td>
</tr>
<tr>
<td>$\sigma_{v77}$</td>
<td>$4.0 \times 10^{-3}$ rad</td>
<td>$4.8 \times 10^{-3}$ rad</td>
</tr>
<tr>
<td>$\sigma_{v88}$</td>
<td>$4.0 \times 10^{-3}$ rad</td>
<td>$4.8 \times 10^{-3}$ rad</td>
</tr>
</tbody>
</table>

Tab. 4.9 Initial parameter values to be used for starting the ML parameter estimation program.
4.6 Simulation Results of the Estimation of the Parameters
in the Flexible Spacecraft Model

Section 4.4 was concerned with the generation of measurement data of a simulated hypothetical flexible spacecraft, which has been discussed in Chapter 2. In this section, the measurement data will be applied to the maximum likelihood parameter estimation of the flexible spacecraft, using the initial estimation of the parameters as discussed in section 4.5.

4.6.1 Simulation results of the estimation of the parameters
using the system model with white process noises

The first experiment for estimating the system parameters was performed, taking into account the following considerations.

1. The measurement data to be used were those from the simulation with the doublet and single pulse input excitations.

2. Two attitude sensors and two attitude rate sensors were used for the rigid main body measurements and four accelerometers were used on the nodes 25, 27, 52 and 54 of the tip nodes for the flexible solar panel measurements as shown in Fig. 4.20.

3. The order of the original model was reduced to 16th by static condensation as discussed in Chapter 3. The translational displacements of nodes 25, 26, 27, 52, 53 and 54 were selected as the master degrees of freedom to be retained in the reduced model.

4. The modeling errors were considered as white noise processes with zero mean and a Gaussian distribution. The number of process noises was 8 for the roll axis, the pitch axis, and the nodes 25, 26, 27, 52, 53 and 54.

From these considerations, the model to be used was written as:

$$\Phi_\mathbf{e} \mathbf{x}(t) = \Phi_\mathbf{e} \mathbf{x}(t) + \Phi_\mathbf{u} \mathbf{u}(t) + \Phi_\mathbf{w} \mathbf{w}(t). \quad (4.39)$$

The state vector, the deterministic input vector and the noise input vector were the following forms:
\[ x = [\phi, \theta, d_{25}^*, d_{26}^*, d_{27}^*, d_{52}^*, d_{53}^*, d_{54}^*] \]

\[ \phi, \theta, d_{25}^*, d_{26}^*, d_{27}^*, d_{52}^*, d_{53}^*, d_{54}^*]^T, \quad (4.40) \]

where \(d_k^*\) denotes the translational displacement of the \(k^{th}\) node.

\[ u = [u_1, u_2]^T, \quad (4.41) \]

where \(u_1\) was the input moment of the roll axis of the control body and \(u_2\) was the input moment of the pitch axis of the control body.

\[ w = [w_1, w_2, \ldots, w_8]^T, \quad (4.42) \]

The matrices in Eq. (4.39) were the following forms:

\[ F_{16 \times 16} = \begin{bmatrix} I_{2x2}^T & A_{2 \times 8}^T & 0_{2 \times 8} \\ A_{2 \times 8} & M_{8 \times 8} & 0_{8 \times 8} \\ 0_{8 \times 2} & 0_{8 \times 6} & U_{8} \end{bmatrix}, \quad (4.43) \]

\[ A_{16 \times 16} = \begin{bmatrix} 0_{2 \times 8} & 0_{2 \times 2} & 0_{2 \times 6} \\ 0_{6 \times 8} & 0_{6 \times 2} & -K_{6} \\ U_{8} & 0_{8 \times 2} & 0_{8 \times 6} \end{bmatrix}, \quad (4.44) \]

\[ B_{16 \times 2} = \begin{bmatrix} h_1 & 0 \ 0 & h_1 \end{bmatrix}_{1 \times 14}^T, \quad (4.45) \]

\[ G_{16 \times 8} = [U_{8}, I_{8}]^T. \quad (4.46) \]

From the consideration 2, the observation equation was be written as:
\[ y(t) = H \cdot x(t) + D \cdot u(t) + Y(t). \quad (4.47) \]

Since eight measurements were used, \( Y(t) \) contained eight components, i.e.:

\[ Y = [y_1 \ y_2 \ \cdots \ y_8]^T. \quad (4.48) \]

From Eq. (3.40), the observation matrix \( H \) was obtained as:

\[ H = (C_1 \ F_r^{-1} \ A_r + C_2). \quad (4.49) \]

In Eq. (4.49) \( C_1 \) followed Eq. (3.33) in Chapter 3:

\[ C_{1}^{\text{4x8}} = \begin{bmatrix}
\mathbb{0}^{4x8} & \mathbb{0}^{4x8} \\
\mathbb{0}^{4x8} & \mathbb{0}^{4x8}
\end{bmatrix}. \quad (4.50) \]

where:

\[ C_{2}^{4x8} = \begin{bmatrix}
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}. \]

The matrix \( C_2 \) in Eq. (4.49) followed Eq. (3.37) in Chapter 3:

\[ C_{2} = \begin{bmatrix}
\mathbb{0}^{4x2} & \mathbb{0}^{4x6} & \mathbb{0}^{4x2} & \mathbb{0}^{4x6} \\
\mathbb{0}^{2x6} & \mathbb{0}^{2x6} & \mathbb{0}^{2x6} & \mathbb{0}^{2x6} \\
\mathbb{0}^{2x6} & \mathbb{0}^{2x6} \ & \mathbb{0}^{2x6}
\end{bmatrix}. \quad (4.51) \]

The feedforward matrix \( D \) was the following form from Eq. (3.41):
\[ D = C_r E_r^{-1} B_r , \] (4.52)

and the measurement noise vector was:

\[ y = C_r E_r^{-1} G w + v_1 , \] (4.53)

where \( w \) is a white process noise vector and \( v_1 \) is a white measurement noise vector.

The parameters to be estimated in this simulated experiment were selected as follows:

1. the diagonal elements of the covariance matrix of the process noises, i.e. \( \sigma_{w_{11}}^2, \sigma_{w_{22}}^2, \ldots, \sigma_{w_{88}}^2 \),
2. the diagonal elements of the covariance matrix of the measurement noises, i.e. \( \sigma_{v_{11}}^2, \sigma_{v_{22}}^2, \ldots, \sigma_{v_{88}}^2 \),
3. the moment of inertia about the roll axis of the rigid control body, i.e. \( I_x \),
4. the moment of inertia about the pitch axis of the rigid control body, i.e. \( I_y \),
5. the Young's modulus \( E \) of the flexible solar panels,
6. the Poisson ratio \( \nu \) of the flexible solar panels, and
7. the input levels of the actuators \( h \) (the levels of the moments on the roll axis and pitch axis of the control body were assumed to be the same).

The mass density \( \rho \) and the dimensions of the elements \( a, b, t \) were assumed to be known since it could be determined during the design of the spacecraft. Therefore \( \rho, a, b \) and \( t \) were chosen as the true values as used in the simulation. The total number of parameters then was 21 due to these discussions.

In Section 4.3 it could be seen that the number of parameters to be estimated in the algorithm solving the Riccati and Lyapunov equations was different from the number of parameters to be estimated in the algorithm without solving the Riccati and Lyapunov equations. Here the number of
parameters to be estimated in each algorithm is examined for the flexible spacecraft model, as described by Eqs. (4.39) and (4.47).
In the algorithm solving the Riccati and Lyapunov equations, it was assumed that the covariance matrices of the process noises and the measurement noises were diagonal matrices. Then each matrix only had eight parameters, viz. the diagonal elements to be estimated. Including the structural parameters, the total number of parameter was 21, as discussed above.
In the algorithm without solving the Riccati and Lyapunov equations, since the acceleration measurements were used, both matrices $K_p$ and $K$ had to be estimated. Considering the system, as described by Eqs. (4.39) and (4.47), the dimensions of $K_p$ and $K$ were $16 \times 8$. Since there was no knowledge about $K_p$ and $K$ before hand, all the elements of $K_p$ and $K$ had to be estimated.
Including the structural parameters, the total number of parameters in the algorithm without solving the Riccati and Lyapunov equations rose to 261.
In the algorithm solving the Riccati and Lyapunov equations, the Kalman gain matrix $K$ can be obtained from the solution of the Riccati equations. The one stage prediction gain matrix $K_p$ can be obtained from the covariance matrices $V_w$ and $V_v$. Furthermore, the matrices $V_w$ and $V_v$ are at least symmetric matrices, if the general case of process and measurement noises is considered. If these noises are independent white noises, the matrices $V_w$ and $V_v$ will be two diagonal matrices. Moreover, the dimensions of the process noise vector $w$ and the measurement noise vector $v$, in most cases, will be lower than the dimension of the state equation. The number of unknown parameters in the algorithms solving the Riccati and Lyapunov equations, therefore, will be less than that of the unknown parameters in the algorithm without solving the Riccati and Lyapunov equations. If too many parameters have to be estimated from a finite number of measurements, the problem of identifiability of all parameters will increase. Because of this consideration, the algorithms solving the Riccati and Lyapunov equations were used to estimate the 21 parameters, rather than estimating the 261 parameters, using the algorithm without solving the Riccati and Lyapunov equations.
The starting values for the diagonal elements of the covariance matrix of assumed process noises can be chosen arbitrarily before running the estimation program. They can be changed according to the convergence rate of the algorithm. When the selected initial estimates of the diagonal elements of the covariance matrix of the process noises are not satisfactory during running the program, new estimates of those values should be used and then the program should be re-started.

In the experiment, the values of the diagonal elements of the covariance matrix of the process noises were chosen as:

\[ V_w = \text{diag.} \{68, 68, 120, 120, 120, 120, 120, 120\} \]  \hspace{1cm} (4.54)

The fit of the estimated state to the eight observed variables after the negative logarithm of the likelihood function became stationary are shown in Figs. 4.41 till 4.57. The results from the runs with doublet and single pulse inputs are nearly the same and given in Tabs. 4.10 till 4.12.

<table>
<thead>
<tr>
<th>parameter</th>
<th>true to be estimated</th>
<th>initial estimates</th>
<th>ML estimates</th>
<th>absolute values of errors</th>
<th>Cramer-Rao lower bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>( I_x )</td>
<td>4480.0</td>
<td>4400.0</td>
<td>4468.71</td>
<td>11.29</td>
<td>5.88</td>
</tr>
<tr>
<td>( I_y )</td>
<td>4480.0</td>
<td>4400.0</td>
<td>4420.58</td>
<td>59.42</td>
<td>106.31</td>
</tr>
<tr>
<td>( E )</td>
<td>( 5.0 \times 10^{10} )</td>
<td>( 6.0 \times 10^{10} )</td>
<td>( 4.65 \times 10^{10} )</td>
<td>( 3.5 \times 10^9 )</td>
<td>( 8.25 \times 10^9 )</td>
</tr>
<tr>
<td>( v )</td>
<td>0.3</td>
<td>0.4</td>
<td>0.34</td>
<td>0.04</td>
<td>0.17</td>
</tr>
<tr>
<td>( h )</td>
<td>200.0</td>
<td>220.0</td>
<td>214.80</td>
<td>14.80</td>
<td>6.62</td>
</tr>
</tbody>
</table>

Tab. 4.10 Structural parameters estimates, errors and Cramer-Rao lower bound of the estimation errors.
<table>
<thead>
<tr>
<th>Parameters to be estimated</th>
<th>Initial estimated</th>
<th>ML estimates</th>
<th>Cramer-Rao lower bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma_{w_{11}}^2$</td>
<td>68.00</td>
<td>15.17</td>
<td>7.51</td>
</tr>
<tr>
<td>$\sigma_{w_{22}}^2$</td>
<td>68.00</td>
<td>81.45</td>
<td>2.43</td>
</tr>
<tr>
<td>$\sigma_{w_{33}}^2$</td>
<td>120.00</td>
<td>69.19</td>
<td>19.15</td>
</tr>
<tr>
<td>$\sigma_{w_{44}}^2$</td>
<td>120.00</td>
<td>12.11</td>
<td>29.18</td>
</tr>
<tr>
<td>$\sigma_{w_{55}}^2$</td>
<td>120.00</td>
<td>81.41</td>
<td>2.45</td>
</tr>
<tr>
<td>$\sigma_{w_{66}}^2$</td>
<td>120.00</td>
<td>72.68</td>
<td>7.01</td>
</tr>
<tr>
<td>$\sigma_{w_{77}}^2$</td>
<td>120.00</td>
<td>7.65</td>
<td>21.33</td>
</tr>
<tr>
<td>$\sigma_{w_{88}}^2$</td>
<td>120.00</td>
<td>39.85</td>
<td>6.11</td>
</tr>
</tbody>
</table>

Tab 4.11 Estimation results variance matrix of process noises $V_{w}$. 
<table>
<thead>
<tr>
<th>Parameter</th>
<th>True Values</th>
<th>Initial Estimates</th>
<th>ML Estimates</th>
<th>Absolute Values of Errors</th>
<th>Cramer-Rao Lower Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma_{v_{11}}^2$</td>
<td>$1.44 \times 10^{-5}$</td>
<td>$9.0 \times 10^{-6}$</td>
<td>$1.407 \times 10^{-5}$</td>
<td>$3.70 \times 10^{-7}$</td>
<td>$6.92 \times 10^{-7}$</td>
</tr>
<tr>
<td>$\sigma_{v_{22}}^2$</td>
<td>$1.44 \times 10^{-5}$</td>
<td>$9.0 \times 10^{-6}$</td>
<td>$1.397 \times 10^{-5}$</td>
<td>$4.70 \times 10^{-7}$</td>
<td>$8.99 \times 10^{-8}$</td>
</tr>
<tr>
<td>$\sigma_{v_{33}}^2$</td>
<td>$1.44 \times 10^{-5}$</td>
<td>$9.0 \times 10^{-6}$</td>
<td>$1.692 \times 10^{-5}$</td>
<td>$2.48 \times 10^{-6}$</td>
<td>$1.15 \times 10^{-5}$</td>
</tr>
<tr>
<td>$\sigma_{v_{44}}^2$</td>
<td>$1.44 \times 10^{-5}$</td>
<td>$9.0 \times 10^{-6}$</td>
<td>$1.450 \times 10^{-5}$</td>
<td>$6.00 \times 10^{-8}$</td>
<td>$3.66 \times 10^{8}$</td>
</tr>
<tr>
<td>$\sigma_{v_{55}}^2$</td>
<td>$2.30 \times 10^{-5}$</td>
<td>$1.6 \times 10^{-5}$</td>
<td>$2.198 \times 10^{-5}$</td>
<td>$1.06 \times 10^{-6}$</td>
<td>$5.10 \times 10^{-6}$</td>
</tr>
<tr>
<td>$\sigma_{v_{66}}^2$</td>
<td>$2.30 \times 10^{-5}$</td>
<td>$1.6 \times 10^{-5}$</td>
<td>$2.400 \times 10^{-5}$</td>
<td>$9.60 \times 10^{-7}$</td>
<td>$2.88 \times 10^{-6}$</td>
</tr>
<tr>
<td>$\sigma_{v_{77}}^2$</td>
<td>$2.30 \times 10^{-5}$</td>
<td>$1.6 \times 10^{-5}$</td>
<td>$2.601 \times 10^{-5}$</td>
<td>$2.97 \times 10^{-7}$</td>
<td>$9.79 \times 10^{-8}$</td>
</tr>
<tr>
<td>$\sigma_{v_{88}}^2$</td>
<td>$2.30 \times 10^{-5}$</td>
<td>$1.6 \times 10^{-5}$</td>
<td>$2.533 \times 10^{-5}$</td>
<td>$2.29 \times 10^{-7}$</td>
<td>$3.18 \times 10^{-8}$</td>
</tr>
</tbody>
</table>

Table 4.12: Estimation results variance matrix of measurement noises $\mathbf{v}$. 
Fig. 4.41 Measurement $y_5$ and estimated $\hat{y}_5$ (roll rate, doublet input).

Fig. 4.42 Measurement $y_6$ and estimated $\hat{y}_6$ (pitch rate, doublet input).

Fig. 4.43 Measurement $y_5$ and estimated $\hat{y}_5$ (roll rate, single pulse input).

Fig. 4.44 Measurement $y_6$ and estimated $\hat{y}_6$ (pitch rate, single pulse input).
Fig. 4.45 Measurement $y_7$ and estimated $\hat{y}_7$ (roll angle, doublet input).

Fig. 4.46 Measurement $y_8$ and estimated $\hat{y}_8$ (pitch angle, doublet input).

Fig. 4.47 Measurement $y_7$ and estimated $\hat{y}_7$ (roll angle, single pulse).

Fig. 4.48 Measurement $y_8$ and estimated $\hat{y}_8$ (pitch angle, single pulse).
Fig. 4.49 Measurement $y_1$ and estimated $\hat{y}_1$ (acceleration of node 25, doublet input).

Fig. 4.50 Measurement $y_2$ and estimated $\hat{y}_2$ (acceleration of node 27, doublet input).

Fig. 4.51 Measurement $y_1$ and estimated $\hat{y}_1$ (acceleration of node 25, single pulse).

Fig. 4.52 Measurement $y_2$ and estimated $\hat{y}_2$ (acceleration of node 27, single pulse).
Fig. 4.53 Measurement $y_3$ and estimated $\hat{y}_3$ (acceleration of node 52, doublet input).

Fig. 4.54 Measurement $y_4$ and estimated $\hat{y}_4$ (acceleration of node 54, doublet input).

Fig. 4.55 Measurement $y_3$ and estimated $\hat{y}_3$ (acceleration of node 52, single pulse).

Fig. 4.56 Measurement $y_4$ and estimated $\hat{y}_4$ (acceleration of node 54, single pulse).
As a summary, the conclusion of to this experiment can be described as follows.

In this experiment, the modelling errors of the reduced order, dynamical model of the spacecraft were assumed to be zero mean white noise processes. This error model is very simple to be used. The estimated time responses of model outputs and the simulated measurement response, as given in Figs. 4.41 till 4.56, show that the high order dynamical model, based on the finite element method, can indeed be reduced to a proper lower order model, which is useful in control system design. The estimated covariance matrix of the process noise \( \hat{v}_w \) is used to fit the estimated response of the model outputs to the observed data. Moreover, some problems remain with the fit of the acceleration measurements. From Figs. 4.29 till 4.32 and Figs. 4.37 till 4.40, it can easily be seen, that the true acceleration responses are not corrupted by noise, whereas the estimated state responses shown in Figs. 4.45 till 4.48 and Figs. 4.53 till 4.56 are clearly noise corrupted. This is due to the fact that the modelling error is assumed to be a white noise process. The estimated acceleration responses consist of the re-constructed process noises \( \hat{v}(k) \), defined by Eq. (4.21), and the re-constructed process noises contain the measurement noises. Therefore, the estimated acceleration responses contain the measurement noises. This is certainly not what is expected, and therefore, the assumption of white noise process for the modelling errors is not valid in this case. This result leads to considering the modelling errors as non-white noises (or colored noises), as discussed in Subsection 3.2.3.

4.6.2 Simulation results of the estimation of the parameters using the model with the colored process noises

As discussed in the conclusions of the last subsection, the white noise process leads in noise corrupted model outputs of the accelerations of the flexible solar panels. In this subsection the following experiment is discussed.

In this experiment with the simulated spacecraft, the modelling errors of the reduced order dynamical modell were assumed as, see Eqs. (3.43) and (3.44) in Chapter 3:
\[ \mathbf{z} = \mathbf{E} \mathbf{z} + \mathbf{w}_2, \]  \hspace{1cm} (4.55)

\[ \mathbf{f} = \mathbf{G}_r \mathbf{z} + \mathbf{G}_1 \mathbf{w}_1, \]  \hspace{1cm} (4.56)

In Eqs. (4.55) and (4.56), \( \mathbf{w}_1 \) and \( \mathbf{w}_2 \) were the vectors of white noise processes with zero mean and a Gaussian distribution. The state vector of the shaping filter \( \mathbf{z} \) consists of six components, which were used for the motions of the flexible solar panels, i.e.:

\[ \mathbf{z} = [z_1 \ z_2 \ z_3 \ z_4 \ z_5 \ z_6]^T. \]  \hspace{1cm} (4.57)

It follows that the noise vector \( \mathbf{w}_2 \) was a \( 6 \times 1 \) vector as follows:

\[ \mathbf{w}_2 = [w_{21} \ w_{22} \ w_{23} \ w_{24} \ w_{25} \ w_{26}]^T. \]  \hspace{1cm} (4.58)

The \( \mathbf{w}_1 \) was a two-dimensional vector:

\[ \mathbf{w}_1 = [w_{11} \ w_{12}]^T, \]  \hspace{1cm} (4.59)

which was used to the motions of the rigid control body.

The vector \( \mathbf{f} \) in Eq. (4.56) was, in fact a \( 16 \times 1 \) vector describing the modelling errors of the reduced system model:

\[ \mathbf{f} = [\mathbf{w}_1 \ \mathbf{z} \ \mathbf{0}_8]^T. \]

Therefore, the matrices \( \mathbf{G}_r \) and \( \mathbf{G}_1 \) were the following forms:
\[
G_r = \begin{bmatrix}
0_{2 \times 6} \\
U_{6} \\
0_{8 \times 6}
\end{bmatrix},
\]

\[
G_1 = \begin{bmatrix}
U_{2} \\
0_{14 \times 2}
\end{bmatrix}.
\]

The matrix \( E \) was assumed as an 6x6 diagonal matrix:

\[
E = -\text{diag.} [e_1, e_2, e_3, e_4, e_5, e_6],
\]

where \( e_i \) was the \( i \)th diagonal element of the matrix \( E \) with positive value, and the minus sign was used to guarantee the shaping filter, as described by Eq. (4.55), to be stable.

In principle, the parameters in the matrix \( E \) should be estimated as well, since the diagonal elements of matrix \( E \), i.e. the correlated time constants of the shaping filter Eq. (4.55), may be different from each other. Therefore, in this system model with white and colored process noises, the number of the parameters to be estimated was six more than the first experiment, where the number of the parameters to be estimated was 21. The number of the parameters to be estimated in this experiment was then increased to 27.

Since the modelling errors have been considered as a dynamical system, the original 16th order dynamical model will increase to a 22th order model. The covariance matrix \( W \) in this experiment was still assumed to be a diagonal matrix, therefore the components in vector \( \bar{z} \) were independent. Other conditions were the same as in the first experiment. The dynamical equation used in this case, were then written as:

\[
F_a x_a(t) = A_a x_a(t) + B_a u(t) + G_a w(t),
\]
where:

\[ x_a = (\phi \theta d_{25}^* d_{26}^* d_{27}^* d_{52}^* d_{53}^* d_{54}^* \phi \theta d_{25}^* d_{26}^* d_{27}^*)^T, \]  \hspace{1cm} (4.64)

\[ u = (u_1 u_2)^T, \]  \hspace{1cm} (4.65)

\[ w = [w_{11} w_{12} w_{21} w_{22} w_{23} w_{24} w_{25} w_{26}]^T, \]  \hspace{1cm} (4.66)

\[ E_{a22x22} = \begin{bmatrix} I & A^T & 0_{2x14} \\ A & M & 0_{6x14} \\ 0_{14x2} & 0_{14x6} & U_{14} \end{bmatrix}, \]  \hspace{1cm} (4.67)

\[ A_{a22x22} = \begin{bmatrix} 0 & 0_{2x6} & 0 & 0_{2x6} & 0_{2x6} \\ 0_{6x2} & 0 & 0_{6x2} & -K & U_6 \\ 0_{2x2} & 0 & 0_{2x6} & 0_{2x2} & 0_{2x6} \\ 0_{6x2} & U_6 & 0_{6x2} & 0_6 & 0_{6x2} \\ 0_{6x2} & 0_6 & 0_{6x2} & 0_6 & 0 \end{bmatrix}, \]  \hspace{1cm} (4.68)

\[ B_{a22x2} = \begin{bmatrix} h & 0 & 0_{2x2}^T \\ 0 & h & 0_{2x2} \end{bmatrix}, \]  \hspace{1cm} (4.69)

\[ G_{a22x8} = \begin{bmatrix} U_{2} & 0_{2x6} \\ 0_{14x2} & 0_{14x6} \\ 0_{6x2} & U_6 \end{bmatrix}. \]  \hspace{1cm} (4.70)

The observation equation from (3.22) was written as:
\[ y(t) = H \ x_a(t) + D \ u(t) + v(t). \]  

(4.71)

Since the number of measurements was still eight, \( y(t) \) is the same as in Eq. (4.49) i.e.:

\[ y = [y_1 \ y_2 \ \ldots \ y_8]^T. \]  

(4.72)

The observation matrix \( H_a \) and the feedforward matrix \( D_a \) in this case were of the forms:

\[ H_a = \begin{pmatrix} C_1 F_1^{-1} & A_r + C_2 F_r^{-1} & G_r \end{pmatrix}, \]  

(4.73)

\[ D_a = C_1 F_1^{-1} B_r. \]  

(4.74)

In Eq. (4.74) \( C_1 \) followed Eq. (3.33) in Chapter 3:

\[ C_1^{-1} = \begin{bmatrix} C_{4 \times 8} & 0_{4 \times 14} \\ 0_{4 \times 8} & 0_{4 \times 14} \end{bmatrix}, \]  

(4.75)

where:

\[ C_{4 \times 8} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \]

The matrix \( C_2 \) in Eq. (4.74) followed Eq. (3.37) in Chapter 3:

\[ C_2 = \begin{bmatrix} 0_{4 \times 2} & 0_{4 \times 6} & 0_{4 \times 2} & 0_{4 \times 12} \\ 0_{2} & 0_{2} & 0_{2} & 0_{2} \\ 0_{2} & 0_{2} & 0_{2} & 0_{2} \end{bmatrix}. \]  

(4.76)
The fit of the estimated responses of the model outputs to the eight observed variables after the likelihood function became stationary, are given in Figs. 4.57 till 4.72. The estimated parameters are shown in Tab. 4.13 till 4.16.

<table>
<thead>
<tr>
<th>parameter true values to be estimated</th>
<th>initial estimates</th>
<th>ML estimates</th>
<th>absolute values of errors</th>
<th>Cramer-Rao lower bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ix</td>
<td>4480.0</td>
<td>4400.0</td>
<td>4498.16</td>
<td>18.16</td>
</tr>
<tr>
<td>Iy</td>
<td>4480.0</td>
<td>4400.0</td>
<td>4329.76</td>
<td>150.24</td>
</tr>
<tr>
<td>E</td>
<td>(5.0 \times 10^{10})</td>
<td>(6.0 \times 10^{10})</td>
<td>(5.28 \times 10^{10})</td>
<td>(2.8 \times 10^9)</td>
</tr>
<tr>
<td>v</td>
<td>0.3</td>
<td>0.4</td>
<td>0.28</td>
<td>0.02</td>
</tr>
<tr>
<td>h</td>
<td>200.0</td>
<td>220.0</td>
<td>216.50</td>
<td>16.50</td>
</tr>
</tbody>
</table>

Tab. 4.13 Estimated structural parameters, errors and the Cramer-Rao lower bound of the estimation errors.

<table>
<thead>
<tr>
<th>Parameters be estimated</th>
<th>initial estimates</th>
<th>ML estimates</th>
<th>Cramer-Rao lower bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>(e_1)</td>
<td>15.00</td>
<td>7.37</td>
<td>2.55</td>
</tr>
<tr>
<td>(e_2)</td>
<td>24.00</td>
<td>2.16</td>
<td>3.97</td>
</tr>
<tr>
<td>(e_3)</td>
<td>15.00</td>
<td>31.45</td>
<td>24.18</td>
</tr>
<tr>
<td>(e_4)</td>
<td>15.00</td>
<td>8.99</td>
<td>5.63</td>
</tr>
<tr>
<td>(e_5)</td>
<td>24.00</td>
<td>7.59</td>
<td>1.01</td>
</tr>
<tr>
<td>(e_6)</td>
<td>15.00</td>
<td>29.68</td>
<td>3.09</td>
</tr>
</tbody>
</table>

Tab. 4.14 Estimated diagonal elements of matrix \(E\) in the shaping filter.
<table>
<thead>
<tr>
<th>Parameters to be estimated</th>
<th>Initial estimates</th>
<th>ML estimates</th>
<th>Cramer-Rao lower bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma_{w11}^2$</td>
<td>68.00</td>
<td>12.51</td>
<td>1.35</td>
</tr>
<tr>
<td>$\sigma_{w22}^2$</td>
<td>68.00</td>
<td>24.67</td>
<td>5.55</td>
</tr>
<tr>
<td>$\sigma_{w33}^2$</td>
<td>80.00</td>
<td>12.99</td>
<td>3.14</td>
</tr>
<tr>
<td>$\sigma_{w44}^2$</td>
<td>80.00</td>
<td>8.19</td>
<td>0.78</td>
</tr>
<tr>
<td>$\sigma_{w55}^2$</td>
<td>80.00</td>
<td>20.15</td>
<td>15.23</td>
</tr>
<tr>
<td>$\sigma_{w66}^2$</td>
<td>80.00</td>
<td>4.88</td>
<td>3.01</td>
</tr>
<tr>
<td>$\sigma_{w77}^2$</td>
<td>80.00</td>
<td>19.16</td>
<td>7.18</td>
</tr>
<tr>
<td>$\sigma_{w88}^2$</td>
<td>80.00</td>
<td>11.80</td>
<td>0.28</td>
</tr>
</tbody>
</table>

Tab 4.15 Estimation results of the diagonal elements of covariance matrix $V_w^2$. 
<table>
<thead>
<tr>
<th>parameter true values to be estimated</th>
<th>initial estimates</th>
<th>ML estimates</th>
<th>absolute values of errors</th>
<th>Cramer-Rao lower bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma_1^2$</td>
<td>$1.444 \times 10^{-5}$</td>
<td>$9.0 \times 10^{-6}$</td>
<td>$1.338 \times 10^{-5}$</td>
<td>$1.06 \times 10^{-6}$</td>
</tr>
<tr>
<td>$\sigma_2^2$</td>
<td>$1.444 \times 10^{-5}$</td>
<td>$9.0 \times 10^{-6}$</td>
<td>$1.291 \times 10^{-5}$</td>
<td>$1.53 \times 10^{-6}$</td>
</tr>
<tr>
<td>$\sigma_3^2$</td>
<td>$1.444 \times 10^{-5}$</td>
<td>$9.0 \times 10^{-6}$</td>
<td>$1.548 \times 10^{-5}$</td>
<td>$1.04 \times 10^{-6}$</td>
</tr>
<tr>
<td>$\sigma_4^2$</td>
<td>$1.444 \times 10^{-5}$</td>
<td>$9.0 \times 10^{-6}$</td>
<td>$1.601 \times 10^{-5}$</td>
<td>$1.57 \times 10^{-6}$</td>
</tr>
<tr>
<td>$\sigma_5^2$</td>
<td>$2.304 \times 10^{-5}$</td>
<td>$1.6 \times 10^{-5}$</td>
<td>$2.211 \times 10^{-5}$</td>
<td>$9.37 \times 10^{-7}$</td>
</tr>
<tr>
<td>$\sigma_6^2$</td>
<td>$2.304 \times 10^{-5}$</td>
<td>$1.6 \times 10^{-5}$</td>
<td>$2.428 \times 10^{-5}$</td>
<td>$1.24 \times 10^{-6}$</td>
</tr>
<tr>
<td>$\sigma_7^2$</td>
<td>$2.304 \times 10^{-5}$</td>
<td>$1.6 \times 10^{-5}$</td>
<td>$2.616 \times 10^{-5}$</td>
<td>$3.12 \times 10^{-6}$</td>
</tr>
<tr>
<td>$\sigma_8^2$</td>
<td>$2.304 \times 10^{-5}$</td>
<td>$1.6 \times 10^{-5}$</td>
<td>$2.100 \times 10^{-5}$</td>
<td>$2.04 \times 10^{-6}$</td>
</tr>
</tbody>
</table>

Tab. 4.16 Estimation results of the diagonal elements of the covariance matrix of measurement noises.
Fig. 4.57 Measurement $y_5$ and estimated $\hat{y}_5$ (roll rate, doublet input).

Fig. 4.58 Measurement $y_6$ and estimated $\hat{y}_5$ (pitch rate, doublet input).

Fig. 4.59 Measurement $y_5$ and estimated $\hat{y}_5$ (roll rate, single pulse).

Fig. 4.60 Measurement $y_6$ and estimated $\hat{y}_6$ (pitch rate, single pulse).
Fig. 4.61 Measurement $y_7$ and estimated $\hat{y}_7$ (roll angle, doublet input).

Fig. 4.62 Measurement $y_8$ and estimated $\hat{y}_8$ (pitch angle, doublet input).

Fig. 4.63 Measurement $y_7$ and estimated $\hat{y}_7$ (roll angle, single pulse).

Fig. 4.64 Measurement $y_8$ and estimated $\hat{y}_8$ (pitch angle, single pulse).
Fig. 4.65 Measurement $y_1$ and estimated $\hat{y}_2$ (acceleration of node 25, doublet input).

Fig. 4.66 Measurement $y_2$ and estimated $\hat{y}_2$ (acceleration of node 27, doublet input).

Fig. 4.67 Measurement $y_1$ and estimated $\hat{y}_1$ (acceleration of node 25, single pulse input).

Fig. 4.68 Measurement $y_2$ and estimated $\hat{y}_2$ (acceleration of node 27, single pulse input).
Fig. 4.69 Measurement $y_3$ and estimated $\hat{y}_3$ (acceleration of node 52, doublet input).

Fig. 4.70 Measurement $y_4$ and estimated $\hat{y}_4$ (acceleration of node 54, doublet input).

Fig. 4.71 Measurement $y_3$ and estimated $\hat{y}_3$ (acceleration of node 52, single pulse input).

Fig. 4.72 Measurement $y_4$ and estimated $\hat{y}_4$ (acceleration of node 54, single pulse input).
In the second experiment, the modelling errors were assumed to be the white noise processes for the motion of the rigid control body and the response of a dynamical system excited by the white noises. The state vector \( \hat{z} \) describing this system extends the original reduced order dynamical model.

It is not necessary to reconstruct the noise processes for the estimated accelerations, since the error vector \( \hat{z} \) has been added as a part of the state vector \( \hat{x} \). Fig. 4.61 till 4.64 and 4.69 till 4.72 show, that the estimated accelerations are indeed smoothed by using the non-white modelling error model.

As a summary to the two experiments in this section, it may be concluded as follows.

1. The white noise process may be used to model the modelling errors resulting from the order reduction, the estimated parameters were reasonably closed to the true parameters in the simulated high order system model. The Cramer-Rao lower bound provided the variances of the errors of the estimated parameters. An advantage of using white noise process is that the order of the reduced order system model remain the same as the original system model. A desvantage of the white process noise is that the reconstructed accelerations of the reduced system model consist of the measurement noises. When the measurement noises are white, the reconstructed accelerations will also be corrupted by the white measurement noises.

2. The modelling errors resulting from the order reduction may assumed to be colored process noises, especially in the case that the accelerations of the reduced model need to be re-constructed. The advantages of using the colored noise for the modelling errors are that the reconstruction of the accelerations is avoided and the reconstructed accelerations are smoothed by using the colored noises. The disadvantages of using the colored noises are that the order of the originally reduced order system model are increased and the number of the parameters to be estimated are also larger than the system model with the white noises processes. As shown in the figures, the fit of the estimated model outputs and the measurements are satisfactory. The estimated parameters are also closed to the parameters in the simulation model.
CHAPTER 5

SUMMARY OF CONCLUSIONS AND SUGGESTIONS FOR FURTHER RESEARCH

5.1 Summary of Conclusions

In this report, the problem of identifying parameters of a flexible spacecraft model from in-orbit measurements was discussed, with emphasis on:

1. mathematical modelling of a flexible spacecraft, using finite element analysis,
2. maximum likelihood parameter estimation, based on system models of reduced order with correlated process and measurement noise.

The main conclusions obtained in this report may be summarized as follows.

1. Finite element analysis can be used to develop mathematical models of finite order of three-dimensional flexible spacecraft in arbitrary orbits.

2. The following parameters in the flexible spacecraft model may be estimated from dynamic response measurements: Young's elasticity modulus, the Poisson ratio, the inertia matrix of the rigid main body and the jet input amplitudes. The remaining parameters are much larger in number, they pertain to the elements of the structural matrices, i.e. the mass, damping and stiffness matrices. The dimension of the parameter estimation problem would rise to an unmanageable level, if these latter parameters were have to be estimated simultaneously with the fewer number of parameters listed first.

3. The required a priori values of the parameters to be estimated from dynamic response measurements, may readily be derived from obvious measurements and experiments prior to launch of the spacecraft.

4. The matrix condensation method used in this report may be used to reduce the order of the original (high order) model, resulting from the finite element analysis. The matrix condensation method generates a quasi-static approximation of the lower frequency characteristic modes. However, the modelling errors which result from such an order reduction have been taken into account. In this study they are treated as unknown
and stochastic disturbances. The time history of these disturbances may be estimated, simultaneously with the unknown parameters listed above.

5. The algorithms derived in Chapter 3 for the computation of maximum likelihood parameter estimation are applicable to the most general case of linear systems with correlated process and measurement noises. Algorithms, which take into account the correlation between process and measurement noises, are essential in the estimation of parameters of flexible spacecraft models of reduced order, if accelerometers are used to measure the dynamics of the spacecraft system to give inputs.

6. An algorithm was derived for the reconstruction of the time histories of unknown noises processes. The algorithm was shown to work successfully in the case, where the external disturbances was a white noise process with a Gaussian distribution.

7. Three versions of the algorithm for the maximum likelihood parameter estimation were derived in Chapter 3, i.e.:
   - the version in which the nonsteady Riccati and Lyapunov equations are solved,
   - the version in which the stable state Riccati and Lyapunov equations are solved and
   - the version in which the solution of the Riccati and Lyapunov equations is avoided.

8. These three versions were applied successfully to the simplified problem of estimating the parameters of a linear second order test model from simulated dynamic response measurements.

9. Subsequently, the maximum likelihood parameter estimation algorithm was applied to a more realistic problem. In this example the version employing the steady state Riccati and Lyapunov equations, was successfully used to estimate the parameters in a linear, reduced order model of a flexible spacecraft with two large flexible solar arrays from simulated response measurements. The modelling errors of the reduced order model were taken into account, assuming unknown white and non-white noise processes.

The objective of this study was to evaluate the feasibility of system identification of flexible spacecraft. Of necessity, the research in this study
was restricted to simulated measurements, and hence the above given conclusions have only a limited scope. However, it is believed that the simulated measurements were sufficiently realistic to admit extrapolation of the conclusions to real, in orbit measurements. Several important questions remain, however, to be answered. The following section gives suggestions for further study in this area.

5.2 Suggestions for further research

The following problems are proposed as examples of further research on the practical application of maximum likelihood parameter identification methods, to flexible spacecraft.

1. Application of algorithms for the calculation of optimal inputs for parameter identification. The performance criterion of such inputs may be some norm of the Cramer-Rao lower bound of the unknown parameters, see e.g. [5.1].

2. Application of the algorithms for maximum likelihood parameter identification, as derived in Chapter 2, to real measurements obtained from a satellite simulator with some flexible appendages.

3. Application of maximum likelihood parameter identification to the case of closed loop attitude control of a flexible spacecraft. This refers to an as yet hypothetical, although perfectly feasible case in which the initial control algorithms are based on a priori model parameter values. Subsequently, the control algorithms are adapted on the basis of parameter identification results.

References

APPENDIX A

THE LAGRANGE'S EQUATIONS FOR QUASI-COORDINATES

The Lagrange's equation of motion, as described in many text books (see, e.g. [2.29] and [2.30]), can be written as the following form:

\[
\frac{d}{dt} \left[ \frac{\partial L}{\partial \dot{q}} \right] - \left[ \frac{\partial L}{\partial q} \right] + \frac{\partial S}{\partial q} = F, \quad \text{(A.1)}
\]

in which \( L \) denotes the so-called Lagrangian, \( q \) and \( \dot{q} \) denote the displacement vector and velocity vector respectively in the inertial reference frame, \( F \) is the external force vector and \( S \) is called the Rayleigh's dissipation function. The Rayleigh's dissipation function has the following form, see [2.29] and [2.30]:

\[
S = \frac{1}{2} \dot{q}^T D \dot{q}, \quad \text{(A.2)}
\]

in which \( D \) is the damping matrix.

The Lagrangian of a general flexible spacecraft may be written as, see Eq. (2.81):

\[
L = \frac{1}{2} \dot{r}^T r \dot{m} + \frac{1}{2} \omega^T \omega + \frac{1}{2} \dot{d}^T M \dot{d} + \dot{r}^T Q \omega + \dot{r}^T W \dot{d} + \omega^T A \dot{d} + \omega^T A \dot{d} + E_{pr}(r) - \frac{1}{2} d^T K d + d^T F_a, \quad \text{(A.3)}
\]

in which the components of \( r \) and \( \omega \) are defined in the rigid body fixed reference frame \( F_b \) and the components of \( d \) are defined in the flexible substructure local reference frames \( F_f \). Since \( F_b \) and \( F_f \) are non-inertial reference frame, \( r, \omega \) and \( d \) are expressed in terms of so-called 'quasi-coordinates', see [2.30].
It can be shown that the Lagrange's equation of motion for this particular case of \( \ell \) being expressed in terms of quasi-coordinates, has the following form, see [2.30]:

\[
\frac{d}{dt}\left\{ \frac{\partial L}{\partial \dot{q}_t} \right\} + \dot{q}_r \left\{ \frac{\partial L}{\partial \dot{q}_t} \right\} - \left\{ \frac{\partial S}{\partial q_t} \right\} + \left\{ \frac{\partial S}{\partial \dot{q}_t} \right\} = F ,
\]

(A.4)

\[
\frac{d}{dt}\left\{ \frac{\partial L}{\partial \dot{q}_r} \right\} + \dot{q}_r \left\{ \frac{\partial L}{\partial \dot{q}_r} \right\} + \dot{q}_t \left\{ \frac{\partial L}{\partial \dot{q}_t} \right\} - \left\{ \frac{\partial S}{\partial q_r} \right\} + \left\{ \frac{\partial S}{\partial \dot{q}_r} \right\} = T ,
\]

(A.5)

in which \( \dot{q}_t \) and \( \dot{q}_r \) denote the velocity and angular velocity vectors, \( q_t \) and \( q_r \) denote the translational and rotational displacement vectors, \( F \) and \( T \) are the external force and torque vectors, and \( \dot{q}_t \) and \( \dot{q}_r \) are skew symmetric matrices of the following form:

\[
g_t = \begin{bmatrix}
0 & -q_{tz} & q_{ty} \\
q_{tz} & 0 & -q_{tx} \\
-q_{ty} & q_{tx} & 0
\end{bmatrix},
\]

\[
g_r = \begin{bmatrix}
0 & -q_{rz} & q_{ry} \\
q_{rz} & 0 & -q_{rx} \\
-q_{ry} & q_{rx} & 0
\end{bmatrix}.
\]

The velocities, angular velocities and displacements in (A.4) and (A.5) are all expressed in terms of quasi-coordinates.

For the rigid main body of the spacecraft, the velocities \( \vec{v} \), the translational displacements \( \vec{r} \) and the angular velocity vector \( \omega \) are defined in the
rigid body fixed reference frame, so that the Lagrange's equation can be written in the following form:

\[
\frac{d}{dt} \left[ \frac{\partial L}{\partial \dot{\omega}} \right] + \omega \left[ \frac{\partial L}{\partial \omega} \right] - \left[ \frac{\partial L}{\partial \dot{r}} \right] = F_b ,
\]

(A.6)

\[
\frac{d}{dt} \left[ \frac{\partial L}{\partial \dot{r}} \right] + \omega \left[ \frac{\partial L}{\partial \omega} \right] + \dot{r} \left[ \frac{\partial L}{\partial \dot{r}} \right] = F_b .
\]

(A.7)

where the skew symmetric matrices \( \omega \) and \( \dot{r} \) are of the following form:

\[
\omega = \begin{bmatrix}
0 & -\omega_z & \omega_y \\
\omega_z & 0 & -\omega_x \\
-\omega_y & \omega_x & 0
\end{bmatrix},
\]

(A.8)

\[
\dot{r} = \begin{bmatrix}
0 & -\dot{r}_z & \dot{r}_y \\
\dot{r}_z & 0 & -\dot{r}_x \\
-\dot{r}_y & \dot{r}_x & 0
\end{bmatrix} .
\]

(A.9)

The elements of \( \omega \) and \( \dot{r} \) are the components of \( \omega \) and \( \dot{r} \), i.e.:

\[
\omega = [\omega_x \ \omega_y \ \omega_z]^T ,
\]

\[
\dot{r} = [\dot{r}_x \ \dot{r}_y \ \dot{r}_z]^T .
\]

In (A.6) and (A.7), describing the rigid body motion, the dissipation function \( S \) is zero, since it is assumed that there is no mass dissipation in the rigid main body during the time involved of interest (constant mass). The
fourth term in (A.5) is also zero since there are no angle variables of the rigid main body in \( L \), see (A.3).

Each term in (A.6) and (A.7) can be obtained for the rigid main body motion as follows:

\[
\frac{d}{dt} \{ \frac{\partial L}{\partial \dot{\gamma}} \} = \frac{d}{dt} \{ m \; \dot{\gamma} + Q \; \dot{\omega} + \hat{W} \; \dot{d} \} = m \; \ddot{\gamma} + Q \; \ddot{\omega} + \hat{W} \; \ddot{d} , \tag{A.10}
\]

\[
\omega \; \{ \frac{\partial L}{\partial \dot{\gamma}} \} = \omega \; \{ m \; \dot{\gamma} + Q \; \dot{\omega} + \hat{W} \; \dot{d} \} , \tag{A.11}
\]

\[- \{ \frac{\partial L}{\partial \epsilon} \} = \frac{\partial E_{pr}}{\partial \epsilon} (r) / \partial \epsilon , \tag{A.12}
\]

\[
\frac{d}{dt} \{ \frac{\partial L}{\partial \omega} \} = \frac{d}{dt} \{ I \; \omega + Q^T \; \dot{r} + \hat{A} \; \dot{d} \} = I \; \ddot{\omega} + Q^T \; \ddot{r} + \hat{A} \; \ddot{d} , \tag{A.13}
\]

\[
\omega \; \{ \frac{\partial L}{\partial \omega} \} = \omega \; \{ I \; \omega + Q^T \; \dot{r} + \hat{A} \; \dot{d} \} , \tag{A.14}
\]

\[
\dot{r} \; \{ \frac{\partial L}{\partial \dot{r}} \} = \dot{r} \; \{ m \; \dot{\gamma} + Q \; \dot{\omega} + \hat{W} \; \dot{d} \} . \tag{A.15}
\]

New, the equations of motion of the rigid main body of the spacecraft can be written as:

\[
m \; \ddot{\gamma} + Q \; \ddot{\omega} + \hat{W} \; \ddot{d} + \omega \; \{ m \; \dddot{\gamma} + Q \; \dddot{\omega} + \hat{W} \; \dddot{d} \} = E_{d} + \frac{\partial E_{pr}}{\partial \gamma} (r) / \partial \gamma , \tag{A.16}
\]

\[
Q^T \; \dddot{r} + I \; \dddot{\omega} + \dot{\hat{A}} \; \dddot{d} + \omega \; \{ Q^T \; \dddot{r} + I \; \dddot{\omega} + \dot{\hat{A}} \; \dddot{d} \} + \dot{r} \; \{ m \; \dddot{\gamma} + Q \; \dddot{\omega} + \hat{W} \; \dddot{d} \} = T_{d} . \tag{A.17}
\]

In (A.16) \( \frac{\partial E_{pr}}{\partial \gamma} (r) / \partial \gamma \) is the external force vector caused by the potential energy of the undeformed spacecraft and \( F_{d} \) is the impressed force vector acting on the rigid main body of the spacecraft. It may be assumed that the
impressed forces, such as the control forces acting on the rigid main body of the spacecraft, are much larger than the forces resulting from the potential energy of the undeformed spacecraft, so that the term $\partial E_{pr}(r)/\partial r$ may be neglected in (A.16). The equation of motion can therefore be written as:

\[ m \ddot{\mathbf{r}} + Q \dot{\omega} + W \ddot{d} + \omega \{ m \ddot{\mathbf{r}} + Q \omega + W \dot{d} \} = F_b \quad , \tag{A.18} \]

\[ Q^T \ddot{\mathbf{r}} + I \dot{\omega} + A \ddot{d} + \omega \{ Q^T \dot{\mathbf{r}} + I \omega + A \dot{d} \} + \dot{r} \{ m \ddot{\mathbf{r}} + Q \omega + W \dot{d} \} = T_b \quad . \tag{A.19} \]

To obtain the equations of motion of the flexible substructures of the spacecraft, each nodal displacement vector $d^*_K$ ($k = 1, 2, \ldots, N_n$) in the total displacement vector $d$ should be arranged in terms of the translational displacements and rotational displacements as follows:

\[ d^*_K = [d^*_kT \quad d^*_kT]^T \quad , \tag{A.20} \]

where each subvector in (A.20) can be written as:

\[ d^*_kT = [d^*_kT_x \quad d^*_kT_y \quad d^*_kT_z]^T \quad , \tag{A.21} \]

\[ d^*_kT = [d^*_kT_x \quad d^*_kT_y \quad d^*_kT_z]^T \quad . \tag{A.22} \]

The components in $d^*_kT$ and $d^*_kT$ are the components of $d^*_kT$ and $d^*_kT$ in the $x$, $y$ and $z$ directions in the flexible substructure local reference frame $F_l$.

The dissipation function due to the structural vibration can be obtained from Rayleigh's dissipation function (A.3) as:
\[ S = \hat{\mathbf{d}}^T \mathbf{D} \hat{\mathbf{d}} \], \hspace{1cm} (A.23)

where the damping matrix \( \mathbf{D} \) has the same dimension as \( \hat{\mathbf{d}} \).

For each node, the equations of motion of the flexible substructures can be obtained as follows:

\[
\frac{d}{dt} \left[ \frac{\partial \mathbf{L}}{\partial \dot{\mathbf{d}}_{\text{kr}}} \right] + \left[ \mathbf{u} \right] \left[ \frac{T_b^T}{\mathbf{d}_{\text{kr}} \mathbf{T}_b} \right] \left[ \frac{\partial \mathbf{L}}{\partial \dot{\mathbf{d}}_{\text{kt}}} \right] + \\
- \left[ \frac{\partial \mathbf{L}}{\partial \dot{\mathbf{d}}_{\text{kt}}} \right] + \left[ \frac{\partial \mathbf{S}}{\partial \dot{\mathbf{d}}_{\text{kt}}} \right] = \mathbf{F}_k , \hspace{1cm} (A.24)
\]

\[
\frac{d}{dt} \left[ \frac{\partial \mathbf{L}}{\partial \dot{\mathbf{d}}_{\text{kr}}} \right] + \left[ \mathbf{u} \right] \left[ \frac{T_b^T}{\mathbf{d}_{\text{kr}} \mathbf{T}_b} \right] \left[ \frac{\partial \mathbf{L}}{\partial \dot{\mathbf{d}}_{\text{kr}}} \right] + \\
+ \left[ \mathbf{u} \right] \left[ \frac{T_b^T}{\mathbf{d}_{\text{kr}} \mathbf{T}_b} \right] \left[ \frac{\partial \mathbf{L}}{\partial \dot{\mathbf{d}}_{\text{kr}}} \right] - \left[ \frac{\partial \mathbf{L}}{\partial \dot{\mathbf{d}}_{\text{kr}}} \right] + \left[ \frac{\partial \mathbf{S}}{\partial \dot{\mathbf{d}}_{\text{kr}}} \right] = \mathbf{T}_k , \hspace{1cm} (A.25)
\]

where \( \mathbf{F}_k \) and \( \mathbf{T}_k \) are the external force and torque vectors acting on the discussed node and the matrices \( \dot{\mathbf{d}}_{\text{kr}} \) and \( \dot{\mathbf{d}}_{\text{kt}} \) are skew symmetric matrices of the following form:

\[
\dot{\mathbf{d}}_{\text{kr}} = \begin{bmatrix}
0 & -\dot{d}_{\text{kr}}^z & \dot{d}_{\text{kr}}^x \\
\dot{d}_{\text{kr}}^z & 0 & -\dot{d}_{\text{kr}}^y \\
-\dot{d}_{\text{kr}}^x & \dot{d}_{\text{kr}}^y & 0
\end{bmatrix} , \hspace{1cm} (A.26)
\]
\[
\mathbf{d}_{kt} = \begin{bmatrix}
0 & -d_{kt_z} & d_{kt_y} \\
d_{kt_z} & 0 & -d_{kt_x} \\
-d_{kt_y} & d_{kt_x} & 0
\end{bmatrix}
\] (A.27)

The matrix \( T_{b,1} \) transforms the elements of \( \mathbf{d}_{kr} \) and \( \mathbf{d}_{kt} \) to the rigid body fixed reference frame \( F_b \), \( \mathbf{d}_{kr} \) and \( \mathbf{d}_{kt} \) being defined in the flexible substructure local reference frame \( F_l \), see (0.5).

The individual terms of (A.24) and (A.25) can be obtained as follows:

\[
\frac{d}{dt}\{\mathbf{aL}/\partial \mathbf{d}_{kt}\} = \frac{d}{dt}\left[ \begin{array}{c}
0^3 \\
\mathbf{U}_3 \\
0^3
\end{array} \right] \left[ \begin{array}{c}
\mathbf{M} \mathbf{d} + \mathbf{W}^T \mathbf{r} + \mathbf{A}^T \omega \\
\mathbf{0}
\end{array} \right]
\]

\[
= \left[ \begin{array}{c}
0^3 \\
\mathbf{U}_3 \\
0^3
\end{array} \right] \left[ \begin{array}{c}
\mathbf{M} \mathbf{d} + \mathbf{W}^T \mathbf{r} + \mathbf{A}^T \omega \\
\mathbf{0}
\end{array} \right] ,
\] (A.28)

\[
\left\{ \mathbf{w} + T_{b,1}^T \mathbf{d}_{kr} T_{b,1} \right\} \frac{d}{dt}\{\mathbf{aL}/\partial \mathbf{d}_{kt}\} =
\]

\[
= \left\{ \mathbf{w} + T_{b,1}^T \mathbf{d}_{kr} T_{b,1} \right\} \left[ \begin{array}{c}
0^3 \\
\mathbf{U}_3 \\
0^3
\end{array} \right] \left[ \begin{array}{c}
\mathbf{M} \mathbf{d} + \mathbf{W}^T \mathbf{r} + \mathbf{A}^T \omega \\
\mathbf{0}
\end{array} \right] ,
\] (A.29)

\[
- \frac{d}{dt}\{\mathbf{aL}/\partial \mathbf{d}_{kt}\} = \left[ \begin{array}{c}
0^3 \\
\mathbf{U}_3 \\
0^3
\end{array} \right] (\mathbf{K} \mathbf{d} - \mathbf{F}_a) ,
\] (A.30)
\[
\{\alpha s/\alpha d_{kt}\} = \begin{bmatrix}
0_3 & U_3 & \ldots & 0_3
\end{bmatrix} \dot{d} \begin{bmatrix}
0_3
\end{bmatrix},
\quad (A.31)
\]

\[
\frac{d}{dt}\{\alpha L/\alpha \dot{d}_{kr}\} = \frac{d}{dt}\begin{bmatrix}
0_3 & U_3 & \ldots & 0_3
\end{bmatrix} \begin{bmatrix}
M \ddot{d} + W^T \dot{r} + A^T \omega
\end{bmatrix}
\]

\[
= \begin{bmatrix}
0_3 & U_3 & \ldots & 0_3
\end{bmatrix} \begin{bmatrix}
M \ddot{d} + W^T \dot{r} + A^T \omega
\end{bmatrix},
\quad (A.32)
\]

\[
\{\dot{\omega} + \bar{T}_{rb,1}^T \alpha_{kr} \bar{T}_{rb,1}\} \{\alpha L/\alpha \dot{d}_{kr}\} =
\]

\[
= \{\dot{\omega} + \bar{T}_{rb,1}^T \alpha_{kr} \bar{T}_{rb,1}\} \begin{bmatrix}
0_3 & U_3 & \ldots & 0_3
\end{bmatrix} \begin{bmatrix}
M \ddot{d} + W^T \dot{r} + A^T \omega
\end{bmatrix},
\quad (A.33)
\]

\[
- \{\alpha L/\alpha \dot{d}_{kr}\} = \begin{bmatrix}
0_3 & U_3 & \ldots & 0_3
\end{bmatrix} (K \ddot{d} - \xi_{a}),
\quad (A.34)
\]

\[
\{\alpha s/\alpha d_{kr}\} = \begin{bmatrix}
0_3 & U_3 & \ldots & 0_3
\end{bmatrix} \dot{d} \begin{bmatrix}
0_3
\end{bmatrix},
\quad (A.35)
\]

\[
\{\ddot{r} + \bar{T}_{rb,1}^T \alpha_{kt} \bar{T}_{rb,1}\} \{\alpha L/\alpha \dot{d}_{kt}\} =
\]

\[
= \{\ddot{r} + \bar{T}_{rb,1}^T \alpha_{kt} \bar{T}_{rb,1}\} \begin{bmatrix}
0_3 & U_3 & \ldots & 0_3
\end{bmatrix} \begin{bmatrix}
M \ddot{d} + W^T \dot{r} + A^T \omega
\end{bmatrix},
\quad (A.36)
\]
Since the above expressions pertain to one node, by collecting the terms of all the nodes of the spacecraft, the equations of motion of the flexible spacecraft can be derived as follows:

\[
\begin{align*}
&\{M \ddot{\omega} + \omega^T \dot{r} + A \omega \dot{\omega} + D \dot{\omega} + K \omega\} + \{(\omega + T_{b,1}^T \dot{d}_k + T_{b,1} \dot{d}) U_{3Nn}^* + \\
&+ [r + T_{b,1}^T \dot{d}_k T_{b,1}] U_{3Nn}\} \{M \dot{\omega} + \omega^T \dot{r} + A \omega \dot{\omega}\} = F_a,
\end{align*}
\]  

(A.37)

where \(U_{3Nn}^*\) and \(U_{3Nn}\) are \(3Nn\) by \(3Nn\) matrices of the following form:

\[
U_{3Nn} = \begin{bmatrix}
U_3 & 0_3 & \cdots & 0_3 \\
& U_3 & 0_3 & \cdots & 0_3 \\
& & \ddots & \ddots & \ddots \\
& & & U_3 & 0_3 \\
& & & & U_3 \\
& & & & & U_3
\end{bmatrix}
\]

and:

\[
U_{3Nn}^* = \begin{bmatrix}
0_3 & U_3 & \cdots & U_3 \\
& 0_3 & U_3 & \cdots & U_3 \\
& & \ddots & \ddots & \ddots \\
& & & 0_3 & U_3 \\
& & & & 0_3 \\
& & & & & U_3
\end{bmatrix}
\]
\[
F_a = \begin{bmatrix} F_1^T & F_2^T & \cdots & F_{k-1}^T & F_k^T & F_{k+1}^T & \cdots & F_{Nn}^T \end{bmatrix}^T.
\]

The equations of motion of the flexible spacecraft can now be written as follows:

\[
m \dddot{r} + Q \ddot{\omega} + W \dddot{d} + \omega \left\{ m \dddot{r} + Q \ddot{\omega} + W \dddot{d} \right\} = F_b, \tag{A.38}
\]

\[
Q^T \dddot{r} + I \dddot{\omega} + A \dddot{d} + \omega \left\{ Q^T \dddot{r} + I \dddot{\omega} + A \dddot{d} \right\} + \omega \left\{ m \dddot{r} + Q \ddot{\omega} + W \dddot{d} \right\} = T_b, \tag{A.39}
\]

\[
\begin{align*}
M \dddot{d} + W^T \dddot{r} + A^T \dddot{\omega} + D \dddot{d} + K \dddot{d} + \{ & \left[ \omega + T_{b,1}^T \right] \dddot{d}_{k} + \left[ \omega + T_{b,1}^T \right] \dddot{d}_{k} \} U_{3Nn}^* + \\
& \left[ \omega + T_{b,1}^T \right] \dddot{d}_{k} + \left[ \omega + T_{b,1}^T \right] \dddot{d}_{k} \} U_{3Nn}^* + \\
& \left[ \omega + T_{b,1}^T \right] \dddot{d}_{k} + \left[ \omega + T_{b,1}^T \right] \dddot{d}_{k} \} U_{3Nn}^* + \\
& \left[ \omega + T_{b,1}^T \right] \dddot{d}_{k} + \left[ \omega + T_{b,1}^T \right] \dddot{d}_{k} \} U_{3Nn}^* + \\
& \left[ \omega + T_{b,1}^T \right] \dddot{d}_{k} + \left[ \omega + T_{b,1}^T \right] \dddot{d}_{k} \} U_{3Nn}^* + \\
& \left[ \omega + T_{b,1}^T \right] \dddot{d}_{k} + \left[ \omega + T_{b,1}^T \right] \dddot{d}_{k} \} U_{3Nn}^* + \\
& \left[ \omega + T_{b,1}^T \right] \dddot{d}_{k} + \left[ \omega + T_{b,1}^T \right] \dddot{d}_{k} \} U_{3Nn}^* + \\
& \left[ \omega + T_{b,1}^T \right] \dddot{d}_{k} + \left[ \omega + T_{b,1}^T \right] \dddot{d}_{k} \} U_{3Nn}^* + \end{align*}
\]

\[
M \dddot{d} + W^T \dddot{r} + A^T \dddot{\omega} = F_a. \tag{A.40}
\]

The equations of motion as described by (A.38), (A.39) and (A.40) are non-linear differential equations, see the fourth terms of (A.38) and (A.39), the fifth term of (A.39) and the sixth term of (A.40).

It may be assumed that for small motions of the rigid main body and the deformations of the flexible substructures, the higher order terms in (A.38), (A.39) and (A.40) can be neglected. The resulting linear equations of motion of the flexible spacecraft are of the following form:

\[
m \dddot{r} + Q \ddot{\omega} + W \dddot{d} = F_b, \tag{A.41}
\]

\[
Q^T \dddot{r} + I \dddot{\omega} + A \dddot{d} = T_b, \tag{A.42}
\]

\[
M \dddot{d} + W^T \dddot{r} + A^T \dddot{\omega} + D \dddot{d} + K \dddot{d} = F_a. \tag{A.43}
\]
In Chapter 2 a linear expression for \( \omega \) was derived, see Eq. (2.6):

\[
\omega = \dot{\theta} + \omega_o \dot{\theta} + \omega_o, \tag{A.44}
\]

where:

\[
\dot{\theta} = [\phi \ \theta \ \psi]^T, \tag{A.45}
\]

\[
\omega_o = [0 \ \omega_o \ 0]^T, \tag{A.46}
\]

and:

\[
\omega_o = \begin{bmatrix}
0 & 0 & -\omega_o \\
0 & 0 & 0 \\
\omega_o & 0 & 0
\end{bmatrix}. \tag{A.47}
\]

With (A.44) the equations of motion as described by (A.41), (A.42) and (A.3) can also be written as:

\[
m \ddot{r} + Q_m (\ddot{\theta} + \omega_o \dot{\theta}) + W \ddot{d} = F_b, \tag{A.48}
\]

\[
\omega_o^T \ddot{r} + \omega_o (\ddot{\theta} + \omega_o \dot{\theta}) + A \ddot{d} = T_b, \tag{A.49}
\]

\[
M \ddot{d} + W^T \ddot{r} + A^T (\ddot{\theta} + \omega_o \dot{\theta}) + D \ddot{d} + K d = F_a. \tag{A.50}
\]

These linearized equations of motion of the flexible spacecraft are used in Chapter 2.
APPENDIX B

The stiffness and mass matrices of a rectangular plate element as derived from the two different shape function matrices \( C \) discussed in Chapter 2 are presented in the present appendix.

The stiffness matrix of a plate rectangular element is a symmetrical matrix of the following form:

\[
K = \frac{E t^3}{12 (1 - \nu^2) ab} \begin{bmatrix}
k_{1,1} & k_{1,2} & \cdots & k_{1,12} \\
k_{2,1} & k_{2,2} & \cdots & k_{2,12} \\
\vdots & \vdots & \ddots & \vdots \\
k_{12,1} & k_{12,2} & \cdots & k_{12,12}
\end{bmatrix}
\]

(B.1)

Where \( E \) is the Young's modulus of the material, \( \nu \) is the Poisson ratio of the material, and \( a, b \) and \( t \) are the width, length and thickness of the element, see Chapter 2.

The following expressions may be derived for the lower triangular elements of this matrix from the shape function matrix (2.94), \( c = \frac{a}{b} \):

\[
\begin{align*}
\k_{1,1} &= 4 (c^2 + c^{-2}) + \frac{1}{5} (14 - 4\nu) \\
\k_{3,1} &= -[2 c^{-2} + \frac{1}{5} (1 + 4\nu)] a \\
\k_{5,1} &= [2 c^{-2} + \frac{1}{5} (1 + 4\nu)] b \\
\k_{7,1} &= -2 (c^2 + c^{-2}) + \frac{1}{5} (14 - 4\nu) \\
\k_{9,1} &= [-c^2 + \frac{1}{5} (1 - \nu)] a \\
\k_{11,1} &= [c^{-2} + \frac{1}{5} (1 - \nu)] b \\
\k_{2,2} &= \left[\frac{4}{3} c^{-2} + \frac{4}{15} (1 - \nu)\right] b^2 \\
\k_{4,2} &= -[2c^{-2} + \frac{1}{5} (1 - \nu)] b \\
\k_{2,1} &= [2 c^{-2} + \frac{1}{5} (1 + 4\nu)] b \\
\k_{4,1} &= 2 (c^2 - 2c^{-2}) - \frac{1}{5} (14 - 4\nu) \\
\k_{6,1} &= [-c^2 + \frac{1}{5} (1 + 4\nu)] a \\
\k_{8,1} &= [c^{-2} - \frac{1}{5} (1 - \nu)] b \\
\k_{10,1} &= -2 (2c^2 - c^{-2}) - \frac{1}{5} (14 - 4\nu) \\
\k_{12,1} &= -[2c^2 + \frac{1}{5} (1 - \nu)] a \\
\k_{3,2} &= -wab \\
\k_{5,2} &= \left[\frac{2}{3} c^{-2} - \frac{1}{15} (1 - \nu)\right] b^2
\end{align*}
\]
\[ k_{6,2} = 0 \]
\[ k_{8,2} = \left[ \frac{1}{3} c^{-2} + \frac{1}{15} (1 - \nu) \right] b^2 \]
\[ k_{10,2} = [-c^{-2} + \frac{1}{5} (1 + 4\nu)] b \]
\[ k_{12,2} = 0 \]
\[ k_{3,3} = \left[ \frac{4}{3} c^2 + \frac{4}{15} (1 - \nu) \right] a^2 \]
\[ k_{5,3} = 0 \]
\[ k_{7,3} = [c^2 - \frac{1}{5} (1 - \nu)] a \]
\[ k_{9,3} = \left[ \frac{1}{3} c^2 + \frac{1}{15} (1 - \nu) \right] a^2 \]
\[ k_{11,3} = 0 \]
\[ k_{4,4} = 4 (c^2 - c^{-2}) + \frac{1}{5} (14 - 4\nu) \]
\[ k_{6,4} = -[2c^2 + \frac{1}{5} (1 + 4\nu)] a \]
\[ k_{8,4} = [-c^{-2} + \frac{1}{5} (1 + 4\nu)] b \]
\[ k_{10,4} = -2(c^2 - c^{-2}) + \frac{1}{5} (14 - 4\nu) \]
\[ k_{12,4} = [-c^{-2} + \frac{1}{5} (1 - \nu)] a \]
\[ k_{5,5} = \left[ \frac{4}{3} c^{-2} + \frac{4}{15} (1 - \nu) \right] b^2 \]
\[ k_{7,5} = [-c^{-2} + \frac{1}{5} (1 + 4\nu)] b \]
\[ k_{9,5} = 0 \]
\[ k_{11,5} = \left[ \frac{1}{3} c^{-2} + \frac{1}{15} (1 - \nu) \right] b^2 \]
\[ k_{6,6} = \left[ \frac{4}{3} c^2 + \frac{4}{15} (1 - \nu) \right] a^2 \]
\[ k_{8,6} = 0 \]
\[ k_{10,6} = [c^2 - \frac{1}{5} (1 - \nu)] a \]
\[ k_{12,6} = \left[ \frac{1}{3} c^2 + \frac{1}{15} (1 - \nu) \right] a^2 \]
\[ k_{7,2} = [-c^{-2} + \frac{1}{5} (1 - \nu)] b \]
\[ k_{9,2} = 0 \]
\[ k_{11,2} = \left[ \frac{2}{3} c^{-2} - \frac{4}{15} (1 - \nu) \right] b^2 \]
\[ k_{4,3} = [-c^{-2} + \frac{1}{5} (1 + 4\nu)] a \]
\[ k_{6,3} = \left[ \frac{2}{3} c^2 - \frac{4}{15} (1 - \nu) \right] a^2 \]
\[ k_{8,3} = 0 \]
\[ k_{10,3} = [2c^2 + \frac{1}{5} (1 - \nu)] a \]
\[ k_{12,3} = \left[ \frac{2}{3} c^2 - \frac{1}{15} (1 - \nu) \right] a^2 \]
\[ k_{5,4} = [-2c^{-2} + \frac{1}{5} (1 + 4\nu)] b \]
\[ k_{7,4} = -2(2c^2 - c^{-2}) - \frac{1}{5} (14 - 4\nu) \]
\[ k_{9,4} = [-2c^2 + \frac{1}{5} (1 - \nu)] a \]
\[ k_{11,4} = [-c^{-2} + \frac{1}{5} (1 - \nu)] b \]
\[ k_{6,5} = vab \]
\[ k_{8,5} = \left[ \frac{2}{3} c^{-2} - \frac{4}{15} (1 - \nu) \right] b^2 \]
\[ k_{10,5} = [c^{-2} - \frac{1}{5} (1 - \nu)] b \]
\[ k_{12,5} = 0 \]
\[ k_{7,6} = [2c^2 + \frac{1}{5} (1 - \nu)] a \]
\[ k_{9,6} = \left[ \frac{2}{3} c^2 - \frac{1}{15} (1 - \nu) \right] a^2 \]
\[ k_{11,6} = 0 \]
\[ k_{7,7} = 4 \left( c^2 + c^{-2} \right) + \frac{1}{5} \left( 14 - 4\nu \right) \]
\[ k_{9,7} = \left[ 2c^2 + \frac{1}{5} \left( 1 + 4\nu \right) \right] a \]
\[ k_{11,7} = \left[ -2c^{-2} + \frac{1}{5} \left( 1 - \nu \right) \right] b \]
\[ k_{8,8} = \left[ \frac{4}{3} c^2 + \frac{4}{15} \left( 1 - \nu \right) \right] b^2 \]
\[ k_{10,8} = \left[ 2c^{-2} + \frac{1}{5} \left( 1 - \nu \right) \right] b \]
\[ k_{12,8} = 0 \]
\[ k_{9,9} = \left[ \frac{4}{3} c^2 + \frac{4}{15} \left( 1 - \nu \right) \right] a^2 \]
\[ k_{11,9} = 0 \]
\[ k_{10,10} = 4 \left( c^2 + c^{-2} \right) + \frac{1}{5} \left( 14 - 4\nu \right) \]
\[ k_{12,10} = \left[ 2c^2 + \frac{1}{5} \left( 1 + 4\nu \right) \right] a \]
\[ k_{11,11} = \left[ \frac{4}{3} c^2 + \frac{4}{15} \left( 1 - \nu \right) \right] b^2 \]
\[ k_{12,12} = \left[ \frac{4}{3} c^2 + \frac{4}{15} \left( 1 - \nu \right) \right] a^2 \]

The so-called compatible shape function matrix of Eq. (2.108) leads to the following expressions for the lower triangular elements of the stiffness matrix:

\[ k_{1,1} = \frac{156}{35} \left( c^2 + c^{-2} \right) + \frac{72}{25} \]
\[ k_{2,1} = \left[ \frac{22}{35} c^2 + \frac{78}{35} c^{-2} + \frac{6}{25} \left( 1 + 5\nu \right) \right] b \]
\[ k_{3,1} = \left[ -\frac{78}{35} c^2 + \frac{22}{35} c^{-2} + \frac{6}{25} \left( 1 + 5\nu \right) \right] a \]
\[ k_{4,1} = \frac{54}{35} c^2 - \frac{156}{35} c^{-2} - \frac{72}{25} \]
\[ k_{5,1} = -\frac{13}{35} c^2 + \frac{78}{35} c^{-2} + \frac{6}{25} \]
\[ k_{6,1} = \left[ -\frac{27}{35} c^2 + \frac{22}{35} c^{-2} + \frac{6}{25} \left( 1 + 5\nu \right) \right] a \]
\[ k_{7,1} = -\frac{54}{35} \left( c^2 + c^{-2} \right) + \frac{72}{25} \]
\[ k_{8,1} = \left( \frac{13}{35} c^2 + \frac{27}{35} c^{-2} - \frac{6}{25} \right) b \]
\[ k_{9,1} = \left( -\frac{27}{35} c^2 - \frac{13}{35} c^{-2} + \frac{6}{25} \right) a \]
\[ k_{10,1} = -\frac{156}{35} c^2 + \frac{54}{35} c^{-2} - \frac{72}{25} \]
\[ k_{11,1} = \left[ -\frac{22}{35} c^2 + \frac{27}{35} c^{-2} - \frac{6}{25} (1 + 5v) \right] b \]
\[ k_{12,1} = \left( -\frac{78}{35} c^2 + \frac{13}{35} c^{-2} - \frac{6}{25} \right) a \]
\[ k_{2,2} = \left( \frac{4}{35} c^2 + \frac{52}{35} c^{-2} + \frac{8}{25} \right) b^2 \]
\[ k_{3,2} = \left[ -\frac{11}{35} (c^2 + c^{-2}) + \frac{1}{50} (1 + 60v) \right] ab \]
\[ k_{4,2} = \left( \frac{13}{35} c^2 - \frac{78}{35} c^{-2} - \frac{6}{25} \right) b \]
\[ k_{5,2} = \left( -\frac{3}{35} c^2 + \frac{26}{35} c^{-2} - \frac{2}{25} \right) b^2 \]
\[ k_{6,2} = \left[ -\frac{13}{70} c^2 + \frac{11}{35} c^{-2} + \frac{1}{50} (1 + 5v) \right] ab \]
\[ k_{7,2} = \left( -\frac{13}{35} c^2 - \frac{27}{35} c^{-2} + \frac{6}{25} \right) b \]
\[ k_{8,2} = \left( \frac{3}{35} c^2 + \frac{9}{35} c^{-2} + \frac{2}{25} \right) b^2 \]
\[ k_{9,2} = \left[ -\frac{13}{70} (c^2 + c^{-2}) + \frac{1}{50} \right] ab \]
\[ k_{10,2} = \left[ -\frac{22}{35} c^2 + \frac{27}{35} c^{-2} - \frac{6}{25} (1 + 5v) \right] b \]
\[ k_{11,2} = \left( -\frac{4}{35} c^2 + \frac{18}{35} c^{-2} - \frac{8}{25} \right) b^2 \]
\[ k_{12,2} = \left[ -\frac{11}{35} c^2 + \frac{13}{70} c^{-2} - \frac{1}{50} (1 + 5v) \right] ab \]
\[ k_{3,3} = \left( \frac{52}{35} c^2 + \frac{4}{35} c^{-2} + \frac{8}{25} \right) a^2 \]
\[ k_{4,3} = \left[ -\frac{27}{35} c^2 + \frac{22}{35} c^{-2} + \frac{6}{25} (1 + 5v) \right] a \]
\[ k_{5,3} = \left[ \frac{13}{70} c^2 - \frac{11}{35} c^{-2} - \frac{1}{50} (1 + 5v) \right] ab \]
\[ k_{6,3} = \left( \frac{18}{35} c^2 - \frac{4}{35} c^{-2} - \frac{8}{25} \right) a^2 \]
\[ k_{7,3} = \left( \frac{27}{35} c^2 + \frac{13}{35} c^{-2} - \frac{6}{25} \right) a \]
\[ k_{8,3} = \left[ \frac{13}{70} (b^2 + b^{-2}) - \frac{1}{50} \right] ab \]
\[ k_{9,3} = \left( \frac{9}{35} \right) c^2 + \frac{3}{35} c^{-2} + \frac{2}{25} \) a^2 \]
\[ k_{10,3} = \left( \frac{78}{35} \right) c^2 - \frac{13}{35} c^{-2} + \frac{6}{25} \) a \]
\[ k_{11,3} = \left[ \frac{11}{35} c^2 - \frac{13}{70} c^{-2} + \frac{1}{50} (1 + 5v) \right] ab \]
\[ k_{12,3} = \left( \frac{26}{35} \right) c^2 - \frac{3}{35} c^{-2} - \frac{2}{25} \) a^2 \]
\[ k_{4,4} = \frac{156}{35} (c^2 + c^{-2}) + \frac{72}{25} \]
\[ k_{5,4} = -\left( \frac{22}{35} c^2 + \frac{78}{35} c^{-2} + \frac{6}{25} (1 + 5v) \right) b \]
\[ k_{6,4} = -\left( \frac{78}{35} c^2 + \frac{22}{35} c^{-2} + \frac{6}{25} (1 + 5v) \right) a \]
\[ k_{7,4} = - \frac{156}{35} c^2 + \frac{54}{35} c^{-2} - \frac{72}{25} \]
\[ k_{8,4} = \left[ \frac{22}{35} c^2 - \frac{21}{35} c^{-2} + \frac{6}{25} (1 + 5v) \right] b \]
\[ k_{9,4} = (\ - \frac{78}{35} c^2 + \frac{13}{35} c^{-2} - \frac{6}{25} ) a \]
\[ k_{10,4} = - \frac{54}{35} (c^2 + c^{-2}) + \frac{72}{25} \]
\[ k_{11,4} = (\ - \frac{13}{35} c^2 - \frac{27}{35} c^{-2} + \frac{6}{25} ) b \]
\[ k_{12,4} = (\ - \frac{27}{35} c^2 - \frac{13}{35} c^{-2} + \frac{6}{25} ) a \]
\[ k_{5,5} = \left[ \frac{4}{35} c^2 + \frac{52}{35} c^{-2} + \frac{1}{50} (1 + 60v) \right] b^2 \]
\[ k_{6,5} = \left[ \frac{11}{35} (c^2 + c^{-2}) + \frac{1}{50} (1 + 60v) \right] ab \]
\[ k_{7,5} = \left[ \frac{22}{35} c^2 - \frac{25}{35} c^{-2} + \frac{6}{25} (1 + 5v) \right] b \]
\[ k_{8,5} = (\ - \frac{4}{35} c^2 + \frac{18}{35} c^{-2} - \frac{8}{25} ) b^2 \]
\[ k_{9,5} = \left[ \frac{11}{35} c^2 - \frac{13}{70} c^{-2} + \frac{1}{50} (1 + 5v) \right] ab \]
\[ k_{10,5} = (\ - \frac{13}{35} c^2 + \frac{27}{35} c^{-2} - \frac{6}{25} ) b \]
\[ k_{11,5} = (\ - \frac{3}{35} c^2 + \frac{9}{35} c^{-2} + \frac{2}{25} ) b^2 \]
\[ k_{12,5} = (\ - \frac{13}{70} (c^2 + c^{-2}) - \frac{1}{50} ) ab \]
\[ k_{6,6} = (\ - \frac{52}{35} c^2 + \frac{4}{35} c^{-2} + \frac{8}{25} ) a^2 \]
\[ k_{7,6} = \left( \frac{78}{35} c^2 - \frac{13}{35} c^{-2} + \frac{8}{25} \right) a^2 \]
\[ k_{8,6} = \left[ - \frac{11}{35} c^2 + \frac{13}{70} c^{-2} - \frac{1}{50} \left( 1 + 5v \right) \right] ab \]
\[ k_{9,6} = \left( \frac{26}{35} c^2 - \frac{3}{35} c^{-2} - \frac{2}{25} \right) a^2 \]
\[ k_{10,6} = \left( \frac{27}{35} c^2 + \frac{13}{35} c^{-2} - \frac{6}{25} \right) a \]
\[ k_{11,6} = \left[ \frac{13}{70} \left( c^2 + c^{-2} \right) - \frac{1}{50} \right] ab \]
\[ k_{12,6} = \left( \frac{9}{35} c^2 + \frac{3}{35} c^{-2} + \frac{2}{25} \right) a^2 \]
\[ k_{7,7} = \frac{156}{35} \left( c^2 + c^{-2} \right) + \frac{72}{25} \]
\[ k_{8,7} = \left[ - \frac{22}{35} c^2 + \frac{78}{35} c^{-2} + \frac{6}{25} \left( 1 + 5v \right) \right] b \]
\[ k_{9,7} = \left[ \frac{78}{35} c^2 + \frac{22}{35} c^{-2} + \frac{6}{25} \left( 1 + 5v \right) \right] a \]
\[ k_{10,7} = \left( \frac{54}{35} c^2 - \frac{156}{35} c^{-2} - \frac{72}{25} \right) b \]
\[ k_{11,7} = \left( \frac{13}{35} c^2 - \frac{78}{35} c^{-2} - \frac{6}{25} \right) b \]
\[ k_{12,7} = \left[ \frac{27}{35} c^2 - \frac{22}{35} c^{-2} - \frac{6}{25} \left( 1 + 5v \right) \right] a \]
\[ k_{8,8} = \left( \frac{4}{35} c^2 + \frac{52}{35} c^{-2} + \frac{8}{25} \right) b^2 \]
\[ k_{9,8} = \left[ - \frac{11}{35} \left( c^2 + c^{-2} \right) + \frac{1}{50} \left( 1 + 60v \right) \right] ab \]
\[ k_{10,8} = \left( - \frac{13}{35} c^2 + \frac{78}{35} c^{-2} + \frac{6}{25} \right) b \]
\[ k_{11,8} = \left( \frac{3}{35} c^2 + \frac{26}{35} c^{-2} - \frac{2}{25} \right) b^2 \]
\[ k_{12,8} = \left[ - \frac{13}{70} c^2 + \frac{11}{35} c^{-2} + \frac{1}{50} \left( 1 + 5v \right) \right] ab \]
\[ k_{9,9} = \left( \frac{52}{35} c^2 + \frac{4}{35} c^{-2} + \frac{8}{25} \right) a^2 \]
\[ k_{10,9} = \left[ \frac{27}{35} c^2 - \frac{22}{35} c^{-2} - \frac{6}{25} \left( 1 + 5v \right) \right] a \]
\[ k_{11,9} = \left( \frac{13}{70} c^2 - \frac{11}{35} c^{-2} - \frac{1}{50} \left( 1 + 5v \right) \right) ab \]
\[ k_{12,9} = \left( \frac{18}{35} c^2 - \frac{4}{35} c^{-2} - \frac{8}{35} \right) a^2 \]
\[ k_{10,10} = \frac{156}{35} \left( c^2 + c^{-2} \right) + \frac{72}{25} \]
\[ k_{11,10} = \left[ \frac{22}{35} c^2 + \frac{78}{35} c^{-2} + \frac{6}{25} (1 + 5v) \right] b \]
\[ k_{12,10} = \left[ \frac{48}{35} c^2 + \frac{22}{35} c^{-2} + \frac{6}{25} (1 + 5v) \right] a \]
\[ k_{11,11} = \left( \frac{4}{35} c^2 + \frac{52}{35} c^{-2} + \frac{8}{25} \right) b^2 \]
\[ k_{12,11} = \left( \frac{11}{35} c^2 + c^{-2} + \frac{1}{50} (1 + 60v) \right) ab \]
\[ k_{12,12} = \left( \frac{52}{35} c^2 + \frac{4}{35} c^{-2} + \frac{8}{25} \right) a^2 \]

The mass matrix of the rectangular plate element is a symmetrical matrix of the following form:

\[
M = \frac{\rho abt}{176,400} \begin{bmatrix}
  m_{1,1} & m_{1,2} & \cdots & m_{1,12} \\
  m_{2,1} & m_{2,2} & \cdots & m_{2,12} \\
  \vdots & \vdots & \ddots & \vdots \\
  m_{12,1} & m_{12,2} & \cdots & m_{12,12}
\end{bmatrix}
\]

For the case of the shape function matrix (2.94) the lower triangular elements are:

\[
m_{1,1} = 24,178 \quad m_{2,1} = 3,227b \quad m_{3,1} = -3,227a \quad m_{4,1} = 8,582 \quad m_{5,1} = -1,918b \\
m_{5,1} = -1,918b \quad m_{6,1} = -1,393a \quad m_{7,1} = 2,758 \quad m_{8,1} = -812b \\
m_{9,1} = 812a \quad m_{10,1} = 8,582 \quad m_{11,1} = 1,393b \quad m_{12,1} = 1,918a \\
m_{2,2} = 560b^2 \quad m_{3,2} = -441ab \quad m_{4,2} = 1,918b \quad m_{5,2} = -420b^2 \\
m_{6,2} = -294ab \quad m_{7,2} = 812b \quad m_{8,2} = -210b^2 \quad m_{9,2} = 196ab \\
m_{10,2} = 1,393b \quad m_{11,2} = 280b^2 \quad m_{12,2} = 294ab \quad m_{6,3} = 280a^2 \\
m_{3,3} = 560b^2 \quad m_{4,3} = -1393a \quad m_{5,3} = 294ab \quad m_{10,3} = -1918a \\
m_{7,3} = -812a \quad m_{8,3} = 196ab \quad m_{9,3} = -210a^2 \quad m_{10,3} = -1918a
\]
\[ m_{11,3} = -294ab \quad m_{12,3} = -420a^2 \]
\[ m_{4,4} = 24,178 \quad m_{5,4} = -3,227b \quad m_{6,4} = -3,227a \quad m_{7,4} = 8,582 \]
\[ m_{8,4} = -1,393b \quad m_{9,4} = 1,918a \quad m_{10,4} = 2,758 \quad m_{11,4} = 812b \]
\[ m_{12,4} = 812a \]
\[ m_{5,5} = 560b^2 \quad m_{6,5} = 441ab \quad m_{7,5} = -1,393b \quad m_{8,5} = 280b^2 \]
\[ m_{9,5} = -294ab \quad m_{10,5} = -812b \quad m_{11,5} = -210b^2 \quad m_{12,5} = -196ab \]
\[ m_{6,6} = 560a^2 \quad m_{7,6} = -1,918a \quad m_{8,6} = 294ab \quad m_{9,6} = -420a^2 \]
\[ m_{10,6} = -812a \quad m_{11,6} = -196ab \quad m_{12,6} = -210a^2 \]
\[ m_{7,7} = 24,178 \quad m_{8,7} = -3,227b \quad m_{9,7} = 3,227a \quad m_{10,7} = 8,582 \]
\[ m_{11,7} = 1,918b \quad m_{12,7} = 1,393a \]
\[ m_{8,8} = 560b^2 \quad m_{9,8} = -441ab \quad m_{10,8} = -1,918b \quad m_{11,8} = -420b^2 \]
\[ m_{12,8} = -294ab \]
\[ m_{9,9} = 560a^2 \quad m_{10,9} = 1,393a \quad m_{11,9} = 294ab \quad m_{12,9} = 280a^2 \]
\[ m_{10,10} = 24,178 \quad m_{11,10} = 3,227b \quad m_{12,10} = 3,227a \]
\[ m_{11,11} = 560b^2 \quad m_{12,11} = 441ab \]
\[ m_{12,12} = 560a^2 \]

Finally, lower triangular elements of the mass matrix as obtained from the compatible shape function matrix are:

\[ m_{1,1} = 24,336 \quad m_{2,1} = 3,432b \quad m_{3,1} = -3,432a \quad m_{4,1} = 8,424 \]
\[ m_{5,1} = -2-28b \quad m_{6,1} = -1,188a \quad m_{7,1} = 2,916 \quad m_{8,1} = -702b \]
\[ m_{9,1} = 702a \quad m_{10,1} = 8,424 \quad m_{11,1} = 1,188b \quad m_{12,1} = 2,028a \]
\[ m_{2,2} = 624b^2 \quad m_{3,2} = -484ab \quad m_{4,2} = 2,028b \quad m_{5,2} = -468b^2 \]
\[ m_{6,2} = -286ab \quad m_{7,2} = 702b \quad m_{8,2} = -162b^2 \quad m_{9,2} = 169ab \]
\[
\begin{align*}
m_{10,2} &= 1.188b \\
m_{11,2} &= 216b^2 \\
m_{12,2} &= 286ab \\
m_3 &= 624a^2 \\
m_4 &= -1.188a \\
m_5 &= 286ab \\
m_6 &= 216a^2 \\
m_7 &= -702a \\
m_8 &= 169ab \\
m_9 &= -162a^2 \\
m_{10} &= -2.028a \\
m_{11,3} &= -286ab \\
m_{12,3} &= -468a^2 \\
m_{4,4} &= 24.336 \\
m_{5,4} &= -3.432b \\
m_{6,4} &= -3.432a \\
m_{7,4} &= 8.424 \\
m_{8,4} &= -1.188b \\
m_{9,4} &= 2.028a \\
m_{10,4} &= 2.916 \\
m_{11,4} &= 702b \\
m_{12,4} &= 702a \\
m_{5,5} &= 624b^2 \\
m_{6,5} &= 484ab \\
m_{7,5} &= -1.188b \\
m_{8,5} &= 216b^2 \\
m_{9,5} &= -286ab \\
m_{10,5} &= -702b \\
m_{11,5} &= -162b^2 \\
m_{12,5} &= -162ab \\
m_{6,6} &= 624a^2 \\
m_{7,6} &= -2.028a \\
m_{8,6} &= 286ab \\
m_{9,6} &= -468a^2 \\
m_{10,6} &= -702a \\
m_{11,6} &= -162ab \\
m_{12,6} &= 162a^2 \\
m_{7,7} &= 24.336 \\
m_{8,7} &= -3.432b \\
m_{9,7} &= 3.432a \\
m_{10,7} &= 8.424 \\
m_{11,7} &= 2.028b \\
m_{12,7} &= 1.188a \\
m_{8,8} &= 624b^2 \\
m_{9,8} &= -484ab \\
m_{10,8} &= -2.028b \\
m_{11,8} &= -468b^2 \\
m_{12,8} &= -286ab \\
m_{9,9} &= 624a^2 \\
m_{10,9} &= 1.188a \\
m_{11,9} &= 286ab \\
m_{12,9} &= 216a^2 \\
m_{10,10} &= 24.336 \\
m_{11,10} &= 3.432b \\
m_{12,10} &= 3.432a \\
m_{11,11} &= 624b^2 \\
m_{12,11} &= 484ab \\
m_{12,12} &= 624a^2
\end{align*}
\]
APPENDIX C

THE FISHER INFORMATION MATRIX OF A DETERMINISTIC SYSTEM

As mentioned in Section 3.4, the Fisher information matrix of a deterministic system can be exactly evaluated in contrast to the situation for the stochastic system (see Section 3.4). This can be proved as follows.

Let the deterministic system be described by:

\[ \mathbf{x}(k+1) = \mathbf{Φ} \mathbf{x}(k) + \Gamma_u \mathbf{u}(k) \quad \tag{C.1} \]

\[ \mathbf{x}(0) = \mathbf{x}_0 \]

\[ \mathbf{y}(k) = \mathbf{Η} \mathbf{x}(k) + \mathbf{D} \mathbf{u}(k) + \mathbf{w}(k) \quad \tag{C.2} \]

As compared to the stochastic system described by Eqs. (3.73) and (3.74), the deterministic system has no stochastic inputs in Eq. (C.1).

The statistics of the observation noise vector \( \mathbf{y} \) are given by:

\[ \text{E} \{ \mathbf{y}(i) \} = \mathbf{0} \quad \tag{C.3} \]

\[ \text{E} \{ \mathbf{y}(i) \mathbf{y}^T(j) \} = \mathbf{V}_w \delta_{i,j} \quad \tag{C.4} \]

The parameters to be identified in the deterministic system are the unknown parameters in the matrices \( \mathbf{Φ}, \; \Gamma_u, \; \mathbf{Η}, \; \mathbf{D}, \; \mathbf{V}_w \) and the initial condition vector \( \mathbf{x}_0 \), i.e.:

\[ \mathbf{θ}_a = [\mathbf{θ}^T \; \mathbf{x}_0^T \; \mathbf{n}^T]^T \quad \tag{C.5} \]

where \( \mathbf{θ} \) contains all the unknown parameters in matrices \( \mathbf{Φ}, \; \Gamma, \; \mathbf{Η}, \; \mathbf{D}, \) and \( \mathbf{n} \) contains the upper-triangular elements or their square roots of the covariance matrix of the observation noises \( \mathbf{V}_w \).

The negative logarithm of the likelihood function of the deterministic system observations described by (C.1) and (C.2) can be written as:
\[ L(\theta, x_0, u) = \frac{N}{2} \ln \det (\Sigma_v) + \frac{1}{2} \sum_{k=1}^{N} \tilde{\gamma}(k) \Sigma_v^{-1} \tilde{\gamma}(k), \]  

(C.6)

where:

\[ \tilde{\gamma}(k) = \gamma(k) - H x(k) - D u(k). \]  

(C.7)

The first order gradient of the likelihood function with respect to \( i \)th element of the parameter vector \( \theta \), i.e., \( \theta_i \), can be obtained from (C.6) as:

\[ \frac{\partial}{\partial \theta_i} [L(\theta, x_0, u)] = \sum_{k=1}^{N} \frac{\partial}{\partial \theta_i} [\tilde{\gamma}(k)] \Sigma_v^{-1} \tilde{\gamma}(k), \]  

(C.8)

where:

\[ \frac{\partial}{\partial \theta_i} [\tilde{\gamma}(k)] = - \left\{ \frac{\partial}{\partial \theta_i} [H] x(k) + H \frac{\partial}{\partial \theta_i} [x(k)] + \frac{\partial}{\partial \theta_i} (D) u(k) \right\}, \]  

(C.9)

\[ \frac{\partial}{\partial \theta_i} [x(k+1)] = \frac{\partial}{\partial \theta_i} (\phi) x(k) + \phi \frac{\partial}{\partial \theta_i} [x(k)] + \frac{\partial}{\partial \theta_i} (\Sigma_u) u(k). \]  

(C.10)

The first order gradient of the likelihood function with respect to the \( i \)th element of the initial condition vector \( x_0 \), i.e., \( x_{0_i} \), can be obtained from (B-6) as:

\[ \frac{\partial}{\partial x_{0_i}} [L(\theta, x_0, u)] = \sum_{k=1}^{N} \frac{\partial}{\partial x_{0_i}} [\tilde{\gamma}(k)] \Sigma_v^{-1} \tilde{\gamma}(k), \]  

(C.11)

where:

\[ \frac{\partial}{\partial x_{0_i}} [\tilde{\gamma}(k)] = - H \frac{\partial}{\partial x_{0_i}} [x(k)]. \]  

(C.12)
\[
\frac{3}{3x_{0}^{1}1}[x(k+1)] = \Phi \frac{3}{3x_{0}^{1}1}[x(k)] .
\]  
(C.13)

The first order gradient of the likelihood function with respect to the \(i\)th element of the parameter vector \(\eta\), i.e., \(\eta_{i}\), can be obtained from (C.6) as:

\[
\frac{3}{3\eta_{1}1}[L(\theta, x_{0}, \eta)] = \frac{N}{2} \text{tr} \left[ \frac{3}{3x_{0}^{1}v} \frac{3}{3\eta_{1}1}(x_{0}v) \right] + \frac{1}{2} \sum_{k=1}^{N} \sum_{v} \bar{z}_{k}(T) \frac{3}{3\eta_{1}1}(x_{0}v) \bar{z}(k) .
\]  
(C.14)

In order to obtain the Fisher information matrix, the second order gradients of the likelihood function with respect to the parameter vectors \(\theta\), \(x_{0}\) and \(\eta\) should be evaluated first. The evaluation is given as follows.

The Fisher information matrix can be written as the following form:

\[
R = \begin{bmatrix}
R_{\theta \theta} & R_{\theta x_{0}} & R_{\theta \eta} \\
R_{x_{0} \theta} & R_{x_{0} x_{0}} & R_{x_{0} \eta} \\
R_{\eta \theta} & R_{\eta x_{0}} & R_{\eta \eta}
\end{bmatrix}
\]  
(C.15)

where:

\[
R_{\theta \theta} = E \left\{ \frac{3}{3\theta \theta}^{T}[L(\theta, x_{0}, \eta)] \right\} ,
\]  
(C.16)

\[
R_{\theta x_{0}} = E \left\{ \frac{3}{3\theta x_{0}}^{T}[L(\theta, x_{0}, \eta)] \right\} = R_{x_{0} \theta}^{T} ,
\]  
(C.17)

\[
R_{\theta \eta} = E \left\{ \frac{3}{3\theta \eta}^{T}[L(\theta, x_{0}, \eta)] \right\} = R_{\eta \theta}^{T} ,
\]  
(C.18)
\[ R^*_{x_0 x_0} = E \left\{ \frac{\partial^2}{\partial x_0 \partial x_0} [L(\theta, x_0, \eta)] \right\} , \] 

(C.19)

\[ R^*_{x_0 \eta} = E \left\{ \frac{\partial^2}{\partial x_0 \partial \eta} [L(\theta, x_0, \eta)] \right\} = R^T_{\eta x_0} , \] 

(C.20)

\[ R^{*\eta\eta} = E \left\{ \frac{\partial^2}{\partial \eta \partial \eta} [L(\theta, x_0, \eta)] \right\} . \] 

(C.21)

The second order derivative of the likelihood function with respect to the \(i^{th}\) and \(j^{th}\) elements of the parameter vector \(\theta\) can be obtained from (C.8) to (C.14) as follows:

\[ \frac{\partial^2}{\partial \theta_i \partial \theta_j} [L(\theta, x_0, \eta)] = \sum_{k=1}^{N} \left\{ \frac{\partial^2}{\partial \theta_i \partial \theta_j} [\tilde{y}^T(k)] \frac{1}{\tilde{y}(k)} \right\} , \] 

(C.22)

where:

\[ \frac{\partial^2}{\partial \theta_i \partial \theta_j} [\tilde{y}(k)] = - \left\{ \frac{\partial^2}{\partial \theta_i \partial \theta_j} [H] \frac{3}{\hat{x}(k)} \hat{x}(k) \right\} + \frac{\partial}{\partial \theta_i} [H] \frac{\partial}{\partial \theta_j} [\hat{x}(k)] \] 

\[ + \frac{\partial}{\partial \theta_j} [H] \frac{\partial}{\partial \theta_i} [\hat{x}(k)] + \frac{\partial^2}{\partial \theta_i \partial \theta_j} [\hat{x}(k)] \] 

\[ + \frac{\partial^2}{\partial \theta_i \partial \theta_j} (\eta) u(k) \] 

(C.23)
\[ + \frac{\partial}{\partial \theta_j} \left( \phi \right) \frac{\partial^2}{\partial \theta^2} \left[ x(k) \right] + \phi \frac{\partial^2}{\partial \theta_i \partial \theta_j} \left[ x(k) \right] + \frac{\partial^2}{\partial \theta_i \partial \theta_j} \left( n_k \right) u(k) \]  

(C.24)

The second order derivative of the likelihood function with respect to the \( i \)th element of the parameter vector \( \theta \) and the \( j \)th element of the initial condition vector \( x_0 \) are obtained as:

\[ \frac{\partial^2}{\partial \theta_i \partial x_0^j} \left[ L(\theta, x_0, n) \right] = \sum_{k=1}^{N} \left\{ \frac{\partial^2}{\partial \theta_i \partial x_0^j} \left[ \tilde{y}^T(k) \right] v^{-1} \tilde{y}(k) + \frac{\partial}{\partial \theta_i} \left[ \tilde{y}^T(k) \right] v^{-1} \frac{\partial}{\partial x_0^j} \left[ \tilde{y}(k) \right] \right\}, \]  

(C.25)

where:

\[ \frac{\partial^2}{\partial \theta_i \partial x_0^j} \left[ \tilde{y}(k) \right] = - \left[ \frac{\partial}{\partial \theta_i} \left( L \right) \right] \frac{\partial}{\partial x_0^j} \left[ x(k) \right] + H \frac{\partial}{\partial \theta_i \partial x_0^j} \left[ x(k) \right] \]  

(C.26)

\[ \frac{\partial^2}{\partial \theta_i \partial x_0^j} \left[ x(k+1) \right] = \frac{\partial}{\partial \theta_i} \left( \phi \right) \frac{\partial}{\partial x_0^j} \left[ x(k) \right] + \phi \frac{\partial^2}{\partial \theta_i \partial x_0^j} \left[ x(k) \right] \]  

(C.27)

The second order derivative of the likelihood function with respect to the \( i \)th element of the parameter vector \( \theta \) and the \( j \)th element of the parameter vector \( n \) are obtained as:

\[ \frac{\partial^2}{\partial \theta_i \partial n_j} \left[ L(\theta, x_0, n) \right] = \sum_{k=1}^{N} \frac{\partial}{\partial \theta_i} \left[ \tilde{y}^T(k) \right] \frac{\partial}{\partial n_j} \left( v^{-1} \right) \tilde{y}(k) \]  

(C.28)
The second order derivative of the likelihood function with respect to the \(i^{th}\) and \(j^{th}\) elements of the initial condition vector \(x_0\) are obtained as:

\[
\frac{\partial^2}{\partial x_0_i \partial x_0_j} [L(\theta; x_0, \eta)] = \sum_{k=1}^{N} \left[ \frac{\partial}{\partial x_0_i} \frac{\partial}{\partial x_0_j} [\tilde{y}^T(k)] \right] \frac{1}{v} \tilde{y} - \left( \frac{\partial}{\partial x_0_i} [\tilde{y}(k)] \right) + \\
+ \frac{\partial}{\partial x_0_i} [\tilde{y}(k)] \frac{1}{v} \tilde{y} - \left( \frac{\partial}{\partial x_0_i} [\tilde{y}(k)] \right), \tag{C.29}
\]

where:

\[
\frac{\partial^2}{\partial x_0_i \partial x_0_j} [\tilde{y}(k)] = - H \frac{\partial^2}{\partial x_0_i \partial x_0_j} [x(k)], \tag{C.30}
\]

\[
\frac{\partial^2}{\partial x_0_i \partial x_0_j} [\dot{x}(k+1)] = - \frac{\partial^2}{\partial x_0_i \partial x_0_j} [x(k)]. \tag{C.31}
\]

The second order derivative of the likelihood function with respect to the \(i^{th}\) element of the initial condition vector \(x_0\) and the \(j^{th}\) element of the parameter vector \(\eta\) are obtained as:

\[
\frac{\partial^2}{\partial x_0_i \partial \eta_j} [L(\theta; x_0, \eta)] = \sum_{k=1}^{N} \frac{\partial}{\partial x_0_i} [\tilde{y}^T(k)] \frac{\partial}{\partial \eta_j} [v^{-1}] \tilde{y}(k). \tag{C.32}
\]

The second order derivative of the likelihood function with respect to the \(i^{th}\) and \(j^{th}\) elements of the parameter vector \(\eta\) are obtained as:

\[
\frac{\partial^2}{\partial \eta_i \partial \eta_j} [L(\theta; x_0, \eta)] =
\]
\[ \begin{align*}
&= \frac{N}{2} \frac{\partial^2}{\partial \eta_i \partial \eta_j} \left[ \ln \det (\Sigma_v) \right] + \frac{1}{2} \sum_{k=1}^{N} \left( \Sigma_v^T(k) \frac{\partial^2}{\partial \eta_i \partial \eta_j} (\Sigma_v^{-1}) \Sigma_v(k) \right) \\
&= \frac{N}{2} \left[ \text{tr} \left[ \frac{\partial}{\partial \eta_j} (\Sigma_v^{-1}) \frac{\partial}{\partial \eta_i} (\Sigma_v) \right] + \text{tr} \left[ \Sigma_v^{-1} \frac{\partial^2}{\partial \eta_i \partial \eta_j} (\Sigma_v) \right] \right] + \\
&\quad + \frac{1}{2} \sum_{k=1}^{N} \text{tr} \left[ \frac{\partial^2}{\partial \eta_i \partial \eta_j} (\Sigma_v^{-1}) \Sigma_v(k) \Sigma_v^T(k) \right].
\end{align*} \]

(C.33)

It should be noted here that in (C.33) if the parameters \( \eta_i \) and \( \eta_j \) are the triangular elements of the covariance matrix of the observation noises, the second order derivative of this matrix should be a zero matrix. However, if \( \eta_i \) and \( \eta_j \) are the square roots of the triangular elements of the covariance matrix of the observation matrix, the second order derivative of this matrix is not a zero matrix.

In the case where the system dynamic model described by (C.1) is deterministic, the first and second order derivatives of the output error vector \( \tilde{y}(k) \), i.e., \( \frac{\partial}{\partial \theta} [\tilde{y}(k)] \), \( \frac{\partial}{\partial x_0} [\tilde{y}(k)] \), \( \frac{\partial^2}{\partial \eta \partial \theta} [\tilde{y}(k)] \), \( \frac{\partial^2}{\partial \eta \partial x_0} [\tilde{y}(k)] \) and \( \frac{\partial^2}{\partial x_0 \partial x_0} [\tilde{y}(k)] \), are also deterministic, see Eqs. (C.9), (C.10), (C.12), (C.13), (C.23), (C.24), (C.26), (C.27), (C.30) and (C.31). Taking the expectations of (C.22), (C.25), (C.28), (C.29), (C.32) and (C.33) and using the definitions described by (C.3) and (C.4), the submatrices in (C.15) can be obtained as follows:

\[ B_{\theta \theta} = \sum_{k=1}^{N} \frac{\partial}{\partial \theta} [\tilde{y}^T(k)] \Sigma_v^{-1} \frac{\partial}{\partial \theta} [\tilde{y}(k)] \]  

(C.34)

\[ B_{\theta x_0} = \sum_{k=1}^{N} \frac{\partial}{\partial \theta} [\tilde{y}^T(k)] \Sigma_v^{-1} \frac{\partial}{\partial x_0} [\tilde{y}(k)] = B^T_{x_0 \theta} \] 

(C.35)
\[ R_{\theta \eta} = 0 = R^T_{\eta \theta} , \quad \tag{C.36} \]

\[ R_{x_0 x_0} = \sum_{k=1}^{N} \frac{\partial}{\partial x_0} [\tilde{y}^T(k)] v^{-1} \frac{\partial}{\partial x_0} [\tilde{y}(k)] , \quad \tag{C.37} \]

\[ R_{x_0 \eta} = 0 = R^T_{\eta x_0} , \quad \tag{C.38} \]

and the \( i, j \)th element of submatrix \( R_{\eta \eta} \) can be obtained as:

\[ R_{\eta_i \eta_j} = \frac{N}{2} \left[ \text{tr} \left( \frac{\partial}{\partial \eta_j} (v^{-1}) \frac{\partial}{\partial \eta_i} (v) \right) + \text{tr} \left( v^{-1} \frac{\partial^2}{\partial \eta_i \partial \eta_j} (v) \right) + \right. \]

\[ + \text{tr} \left( \frac{\partial^2}{\partial \eta_i \partial \eta_j} (v^{-1}) v \right) \right] . \tag{C.39} \]

In (C.39), the second order derivatives of the inverted covariance matrix of the observation noises with respect to its elements, i.e. \( \eta_i \) and \( \eta_j \), can be written as:

\[ \frac{\partial^2}{\partial \eta_i \partial \eta_j} (v^{-1}) = v^{-1} \frac{\partial}{\partial \eta_j} (v) v^{-1} \frac{\partial}{\partial \eta_i} (v) v^{-1} \frac{\partial}{\partial \eta_j} (v) v^{-1} - v^{-1} \frac{\partial^2}{\partial \eta_i \partial \eta_j} (v) v^{-1} + \]

\[ + v^{-1} \frac{\partial}{\partial \eta_i} (v) v^{-1} \frac{\partial}{\partial \eta_j} (v) v^{-1} \]

\[ = \left[ - \frac{\partial}{\partial \eta_j} (v^{-1}) \frac{\partial}{\partial \eta_i} (v) \right] - \frac{\partial}{\partial \eta_i} (v^{-1}) \frac{\partial}{\partial \eta_j} (v) + \]

\[ - v^{-1} \frac{\partial^2}{\partial \eta_i \partial \eta_j} (v) v^{-1} \right] , \tag{C.40} \]

and therefore:
\[
\frac{\partial^2}{\partial \eta_1 \partial \eta_j} (y_{\eta}^{-1}) y_{\eta} = - \frac{\partial}{\partial \eta_1} (y_{\eta}^{-1}) \frac{\partial}{\partial \eta_j} (y_{\eta}) - \frac{\partial}{\partial \eta_1} (y_{\eta}^{-1}) \frac{\partial}{\partial \eta_j} (y_{\eta}) + \\
- y_{\eta}^{-1} \frac{\partial^2}{\partial \eta_1 \partial \eta_j} (y_{\eta}).
\]  

(C.41)

Substituting (C.41) in (C.39), it follows that:

\[
R_{\eta_1 \eta_j} = \frac{N}{2} \text{tr} \left[ - \frac{\partial}{\partial \eta_1} (y_{\eta}^{-1}) \frac{\partial}{\partial \eta_j} (y_{\eta}) \right]
\]

\[
= \frac{N}{2} \text{tr} \left[ y_{\eta}^{-1} \frac{\partial}{\partial \eta_1} (y_{\eta}) y_{\eta}^{-1} \frac{\partial}{\partial \eta_j} (y_{\eta}) \right].
\]

(C.42)

According to (C.34) to (C.42), the Fisher information matrix of a deterministic system described by (C.1) and (C.2) can now be written as:

\[
R = \begin{bmatrix}
R_{\theta \theta} & R_{\theta x_0} & 0 \\
R_{\theta x_0}^T & R_{x_0 x_0} & 0 \\
0 & 0 & R_{n n}
\end{bmatrix}.
\]

(C.43)

The inverted Fisher information matrix \( R^{-1} \) provides the lower bound of the errors of the estimated parameters, see Section 3.6. Summarizing it can be said, that in order to obtain the lower bound of the errors of the estimated parameters, one should evaluate the sensitivity expressions given by (C.10) and (C.13) for the parameter vector \( \theta \) and the initial condition vector \( x_0 \). The lower bound of the errors of the estimated covariance matrix of the observation noises can be calculated from (C.42).
SAMENVATTING IN HET NEDERLANDS (SUMMARY IN DUTCH)

Bij de parameter-identificatie van flexibele ruimtevaartuigen is een wezenlijk probleem het opstellen van een wiskundig model dat de bewegingen van het ruimtevaartuig voldoende nauwkeurig beschrijft maar waarvan de orde toch niet te groot - en waarin het aantal onbekende parameters beperkt is. Hoe groter de orde van het systeem-model en het aantal onbekende parameters, des te moeilijker is het de parameterwaarden te schatten uit de metingen van de responsie van het ruimtevaartuig.

In dit rapport wordt gebruik gemaakt van de eindige elementenmethode om te komen tot een dynamisch model van het flexibele ruimtevaartuig. Het voordeel van de toepassing van de eindige elementenmethode voor het opstellen van een model is dat de orde van het model eindig is, en dat het aantal onbekende parameters zeer klein kan zijn. Vervolgens wordt de orde van het model teruggebracht tot een gemakkelijker te hanteren niveau door de eigenbewegingen met eigen frequenties boven een vooraf gekozen waarde quasi-statisch te benaderen.

Bij de toepassing van de eindige elementenmethode en de quasi-statische benadering van bepaalde eigenbewegingen moet rekening worden gehouden met het ontstaan van modelfouten. In het rapport wordt voorgesteld de modelfouten te interpreteren als realisaties van een (multi-variabel) stochastisch proces dat werkt op de ingang van het hierboven genoemde vereenvoudigde model (systeem-ruis). Het voordeel van deze aanpak is dat modelfouten nu expliciet mee in beschouwing kunnen worden genomen bij de schatting van de parameter waarden van het vereenvoudigde model.

De 'maximum likelihood' methode is in dit rapport toegepast voor de berekening van parameterwaarden uitgaande van (gesimuleerde) metingen van de responsie van het ruimtevaartuig op bepaalde opzettelijk geïmplementeerde verstoringen. Als gevolg van de gekozen modellerings van de modelfouten blijkt een correlatie tussen de systeemruis en de ruis in de metingen mogelijk te zijn. In het rapport werden meerdere algoritmen ontwikkeld voor de schatting van parameterwaarden waarin met deze correlatie wordt rekening gehouden.
De ontwikkelde algoritmen werden geëvalueerd met behulp van gesimuleerde metingen van de reponsies van een eenvoudig tweede orde systeem en van een satelliet met twee grote flexibele zonnepanelen. In het geval van de satelliet kon worden nagegaan of modelfouten ten gevolge van quasi-statische benaderingen inderdaad kunnen worden behandeld als realisaties van een stochatisch proces.

De goede werking van de ontwikkelde algoritmen kon worden aangetoond aan de hand van de simulaties van het tweede orde systeem. Uit de simulatie-experimenten met de flexibele satelliet bleek dat alle onbekende parameterwaarden in een systeem model van gereduceerde orde goed kunnen worden geschat, en dat het inderdaad mogelijk is modelfouten ten gevolge van quasi-statische benaderingen te interpreteren als realisaties van een stochastisch proces.