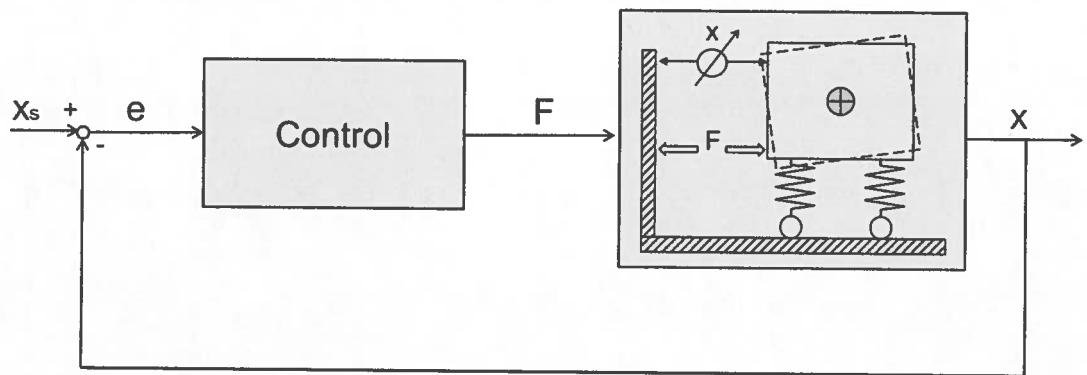


Machine Dynamics in Mechatronic Systems

An Engineering Approach



Adrian M. Rankers

MACHINE DYNAMICS IN MECHATRONIC SYSTEMS
An Engineering Approach

PROEFSCHRIFT

ter verkrijging van
de graad van doctor aan de Universiteit Twente,
op gezag van de rector magnificus,
prof.dr. F.A. van Vught,
volgens besluit van het College voor Promoties
in het openbaar te verdedigen
op vrijdag 13 juni 1997 te 13.15 uur.

door

Adrian Mathias Rankers
geboren op 24 september 1960
te Bad Gandersheim (Duitsland)

Dit proefschrift is goedgekeurd door de promotor

Prof.dr.ir. M.P. Koster

“It is better to be roughly right than precisely wrong”

J.M. Keynes

ISBN: 90-365-0957-2

© Philips Electronics N.V. 1997

Preface

The insights and working methods described in this thesis have been derived during the past 12 years that I have been working on various projects at the Centre for Manufacturing Technology (CFT) of Philips Electronics N.V. I want to acknowledge my gratitude to the management of the CFT for the opportunity of publishing this thesis. I sincerely hope that the quality of the work is such that others will profit from reading it.

Special thanks go to my supervisor, Rien Koster, for his patience, continuous encouragement, and constructive criticism, and to Jan van Eijk, who has been a stimulating coach, teacher, and friend ever since I started working at the CFT.

Furthermore, thanks to all the Philips colleagues and students who have helped me, for rarely is a thesis the work of the author only. Especially I want to thank the members of the project team "Predictive Modelling of Compact Disc Drives"; our team effort back in 1988 was really the starting point for the understanding and predicting of machine dynamics in mechatronic systems. The development in this field of expertise is clearly also boosted by the wafer-stepper activities of ASM-L, in which I was luckily able to participate.

Finally, I want to thank my colleagues of the group "Predictive Modelling" for their co-operation and patience, especially during the last year when I was away so often, working at home on this thesis.

Summary

Designers of products and production equipment nowadays are confronted with increasing demands for higher performance, e.g. more functionality, higher operating speeds and greater precision. Typical examples of this are the well-known Compact Disc Player and equipment that is used in the production of ICs. As a result of these tighter requirements there is a growing trend towards the use of built-in control systems. In the development of these servo-controlled positioning devices it is essential to consider the effect of the dynamics of the mechanical system on the performance of the overall system, because the following effects can be observed :

- Mechanical resonances can endanger the stability of the control loop, and thus limit the maximum bandwidth and the amount of disturbance rejection.
- Vibrations, which are caused by the servo forces during a set-point motion, can lead to positional errors. In a point-to-point application this may introduce the need for extra settling time, or it can even be completely intolerable in a tracking application.

Despite the fact, that mechanical vibrations in a servo device can be very complex and often involve the motion of many components of the system, there are three fundamental mechanisms that are often observed. These three basic dynamic phenomena can best be indicated by :

- Actuator flexibility
- Guiding system flexibility
- Limited mass and stiffness of the stationary machine part

The basic characteristic of what is called actuator flexibility is the fact that the mechanical system does not behave as one rigid body, due to flexibility between the location at which the servo force is applied and the actual point that needs to be positioned. The second dynamic phenomenon results from guiding system flexibility in combination with the fact that the device is driven in such a way that it has to rely on the guiding system to suppress motion in an undesired direction (in case of a linear direct drive system this occurs if the driving force is not applied at the centre of gravity). Due to the limited mass and limited support stiffness of the stationary machine part, the reaction force that comes with the driving force will introduce a motion of the "stationary" part of the mechanical system.

Whereas the first two phenomena mainly affect the stability of the control loop, the last phenomena manifests itself more often as a dynamic positional error in the set-point response.

A tool that can be very useful in understanding the nature of more complex resonance phenomena and the underlying motion of the mechanical system, is “Modal Analysis” or “Modal Decomposition”. This approach already has a long tradition as an experimental and numerical technique. Translating the mathematics of one single decoupled “modal” equation into a graphical representation, which includes all relevant data such as (effective) modal mass and stiffness plus the motion of each physical DOF, facilitates a better understanding of the modal concept. It enables a very intuitive link between the modal and the physical domain, and thus leads to a more creative use of “modal analysis” without the complications of the mathematical formalism.

Dynamic phenomena of the mechanics in a servo positioning device can lead to stability problems of the control loop. Therefore it is important to investigate the frequency response ($x_{\text{servo}}/F_{\text{servo}}$), which characterises the dynamics of the mechanical system, and especially the influence of mechanical resonances on it. Once the behaviour of one individual mode is fully understood it is not so difficult to construct this frequency response and the interaction between the rigid-body motion of the device (leading to a frequency response with a -2 slope), and the dynamics of one additional mode. Four interaction patterns can be distinguished :

- -2 slope/zero/pole/-2 slope
- -2 slope/pole/zero/-2 slope
- -2 slope/pole/-4 slope
- -2 slope/pole/-2 slope

The phase plot corresponding to a -2 slope/zero/pole/-2 slope behaviour displays a phase lead between the frequency of the zero and the pole, whereas the reverse order of pole and zero in the case of a -2 slope/pole/zero/-2 slope pattern leads to a phase lag. The phase of the third characteristic shape, -2 slope/pole/-4 slope, rotates from -180° to -360° at the frequency of the pole. The last pattern, which corresponds to a non-minimum phase system, should generally not occur in a well-designed system, and is therefore excluded from further analysis. The suggested approach for the construction of frequency response functions resulting from one rigid body motion in combination with one additional mode can also be applied very successfully to the three basic dynamic phenomena.

It is not possible to judge the potential destabilising effect of each of the three typical characteristics without considering the frequency of the resonance in relation to the envisaged bandwidth of the control loop. The phase plot of a typical open loop frequency response of a PID controlled positioning device without mechanical

resonances can be divided into three frequency ranges, which are separated by -180° transitions. At low frequencies the phase lies below -180° due to the integrator action of the controller. In the medium-frequency range (centred by the bandwidth frequency) the phase lies above -180° due to the differential action of the controller, which is necessary in order to achieve a stable position control-loop. At high frequencies the phase eventually drops again below -180° due to additional low-pass filtering, limited bandwidth of components in the loop, or limited sampling-rate of a digital controller.

With this ideal open-loop behaviour in mind, the potential destabilising effect of each of the three typical characteristics can be judged in relation to the frequency range (low/medium/high). Whether instability really occurs depends very strongly on the resonance amplitude and damping of the additional mode.

- A -2 slope/zero/pole/-2 slope characteristic leads to a phase lead and is therefore potentially destabilising in the low-frequency and high-frequency regions. In the medium-frequency region it adds an extra phase lead to the already existing margin, which does not harm the stability.
- A -2 slope/pole/zero/-2 slope combination has the reverse effect. It is potentially destabilising in the medium-frequency range and is harmless in the low- and high-frequency ranges.
- The -2 slope/pole/-4 slope behaviour always has a devastating effect on the stability of the loop if located in the low- or medium-frequency range. In the high-frequency range it can be dealt with.

On the basis of these considerations it is possible to give design guidelines for servo positioning devices. These guidelines focus on the three basic dynamic phenomena.

The subject of machine dynamics and its interaction with the control system plays a dominant role in fast and accurate positioning devices, so it is vital to consider these issues during the entire design process. Modelling and simulation can be adequate tools for that purpose; however, two conditions are crucial to the success :

- usefulness of results
- speed

It is apparent, yet often neglected, that analysis results with the relevant amount of detail and accuracy need to be available at the decision-making moment. Furthermore, the analysis data often need to be transformed into useful information on the basis of which design decisions can be motivated.

These requirements can only be achieved by an analysis process which is closely linked to the phases in the design process. During the design process, which generally starts with various conceptual designs, and then converges towards one final detailed design, the amount of information and detail continuously grows.

Consequently, the analysis process should have a similar top-down structure. Starting with very elementary simulation models to support the selection of the proper concept, these models should become more refined, just like the product or machine under development.

In various projects throughout the past years a three-step modelling approach has evolved, in which the following phases can be distinguished :

- concept analysis
- system analysis
- component analysis

In the concept analysis the viability of various concepts is evaluated on the basis of very simple models consisting of a limited number of lumped masses connected by springs. Once a concept has been chosen and the first rough 3D sketches become available, a system analysis can be done, based on a limited number of 3D rigid components connected by springs. In this phase a lot of important spatial information is added to the model (such as the location of the centres of gravity and connecting stiffnesses, plus the location of the driving force and of the sensor). Finally, in the component analysis phase critical components are no longer considered rigid, and their internal dynamics are evaluated via FE modelling. In cases in which a separate analysis of a critical component is considered insufficient to judge its influence on the overall dynamics, a detailed FE-based description can be used to replace the former rigid description in the system model.

In case many parts of the system need to be modelled in great detail, it is not very practical (error-prone, huge model size, time consuming) to build one, single, huge FE model of the entire system. A technique that overcomes these disadvantages is the so-called "Substructuring Technique". In this approach the system is divided into substructures or components, which are analysed separately. Then, after application of a reduction technique which preserves the most dominant dynamic properties, the (reduced) models of the components are assembled to form the overall system. By doing so, the size of the final system model is reduced significantly. The reduction techniques used for substructuring can roughly be divided into two categories :

- Static Reduction Techniques
- Component Mode Techniques

It is argued why the component mode technique based on the so-called Craig-Bampton approach is preferred.

In some cases, time-domain simulations of a manipulator, including its servo controller, cannot be done with the modelling tools used for the analysis of the mechanical system. As a result, the dynamic properties of the mechanics need to be transferred to some other modelling tool which allows to add the control system.

Generally, this transfer is based on a truncated modal description of the mechanical system which is translated into a state-variable representation.

Many of the insights and techniques described in this thesis have evolved while working on projects related to compact disc drives. Development of these devices is a very challenging task, because submicron accuracy has to be achieved at a discount price and under extreme operating conditions (temperature range, jogging, ..). The existence of mechanical resonances and its impact on the control loop is one of the limiting factors in the design of such a device and will probably remain an important issue in the future, because specifications in terms of critical dimensions and speed are ever increasing. Practically all topics discussed in this thesis are covered by the case study of the development of the CDM-8 compact disc module.

Through the enormous performance drive in mechatronics systems much has been learned in the past years about the influence of machine dynamics in servo positioning devices. The lack of complicated mathematics in this thesis shows that these important aspects can be described and understood using operational knowledge of basic machine dynamics and control issues only. Therefore, the content of this thesis is considered to be easily transferable knowledge, which hopefully will be included in the education of mechatronics engineers in the near future.

Contents

PREFACE	i
SUMMARY	iii
1. INTRODUCTION	1
1.1. General	1
1.2. State of the Art	2
1.3. Scope and Purpose	4
1.4. Preview	5
2. MECHANICAL SERVO SYSTEMS	7
2.1. Basic Control Aspects	9
2.2. Specifications	13
2.3. Interaction Dynamics and Control	14
2.4. Three Important Dynamic Effects	16
2.4.1. <i>Actuator Flexibility</i>	16
2.4.2. <i>Guiding System Flexibility</i>	17
2.4.3. <i>Limited Mass and Stiffness of Stationary Machine Part</i>	19
3. MODAL DECOMPOSITION	21
3.1. Mathematics of Modal Decomposition	21
3.2. Graphical Representation	25
3.2.1. <i>Lumped Mass Model</i>	25
3.2.2. <i>General System</i>	28
3.2.3. <i>User-Defined DOF</i>	30
3.3. Physical Meaning of Modal Parameters	32
3.4. A Pragmatic View on Sensitivity Analysis	38
3.5. Modal Superposition	43

3.6. Suspension Modes	45
4. MODES & SERVO STABILITY	49
4.1. Basic Characteristics of Mechanical FRF	49
4.2. Destabilising Effect of Modes	59
4.2.1. <i>FRF of type “-2 slope/zero/pole/-2 slope”</i>	61
4.2.2. <i>FRF of type “-2 slope/pole/zero/-2 slope”</i>	63
4.2.3. <i>FRF of type “-2 slope/pole/-4 slope”</i>	65
4.2.4. <i>Resume</i>	67
4.3. Design for Stability	67
4.3.1. <i>Actuator Flexibility</i>	67
4.3.2. <i>Guiding System Flexibility</i>	73
4.3.3. <i>Limited Mass and Stiffness of Stationary Machine Part</i>	78
4.3.4. <i>General Guidelines</i>	80
5. PREDICTIVE MODELLING	85
5.1. Steps in a Modelling Activity	85
5.2. Stepwise Refined Modelling	88
5.3. Practical Modelling Issues	95
5.3.1. <i>Mixed Dynamics/Control Simulations</i>	99
5.3.2. <i>Effect of Modal Truncation and Accuracy</i>	101
5.3.1 <i>Model Reduction Techniques</i>	109
5.3.1.1. <i>Static Reduction Techniques</i>	110
5.3.1.2. <i>Component Mode Techniques</i>	111
6. CASE STUDY OF A CD-DRIVE	115
6.1. Basics of a Compact Disc System	116
6.2. Servo Specifications	117
6.3. Concept Evaluation	119
6.4. System Evaluation	126
6.4.1. <i>Preliminary Actuator Model</i>	127
6.4.2. <i>Overall Model</i>	136
6.4.3. <i>Slide Servo System</i>	138
6.4.4. <i>Radial Servo System</i>	139
6.4.5. <i>Focus Servo System</i>	142

6.5. Component Evaluation	147
7. CONCLUSIONS	151
REFERENCES	153
NOTATION	157
APPENDIX	159
A. Effect of Frame Dynamics on Settling Time	159
B. Static Reduction Techniques	168
C. Component Modes	172
D. Extra Reduction of Model Size	175
E. Substructure Coupling	179
F. User-Defined DOF	181
SAMENVATTING	185
CURRICULUM VITAE	191

1. Introduction

1.1. General

Designers of products and production equipment nowadays are confronted with increasing demands for higher performance, e.g. more flexibility, higher operating speeds and greater precision (typical examples of this drive are the well-known Compact Disc Player, or equipment that is used in the production of ICs). As a result of these requirements there is a growing trend towards the use of built-in control systems. In the development of these servo-controlled positioning devices it is essential to consider the effect of the dynamics of the mechanical system on the performance of the overall, because the following effects can be observed :

- Mechanical resonances can endanger the stability of the control loop and thus limit its performance.
- Vibrations of the mechanical system, which are caused by the servo forces during a prescribed motion, can lead to positional errors. In point-to-point applications this can introduce the need for extra settling time such that the positional error can decline, whereas in tracking or contouring applications the positional error must remain below a certain threshold at all times.

To obtain a well-balanced design with respect to the effort in the mechanical design and the control design, one has to adapt a mechatronics approach in which the structural design and the control design are integrated. This requires an operational understanding of machine dynamics and basic control systems, and of the interaction between these two disciplines. Furthermore, integrated modelling and simulation of structural and control aspects should be part of the product-creation process of any mechatronic positioning device from the very beginning. Such an approach is the only way to enhance the score of success and achieve "First-time-right".

1.2. State of the Art

The topics of control theory and machine dynamics already have a long history, and a large amount of theoretical knowledge has been built up. Also, a growing number of numerical and experimental tools, such as modal testing¹ and FE-analysis, is available. Still, one can observe that in the design of machines and products only limited use is made of these tools and the available expertise.

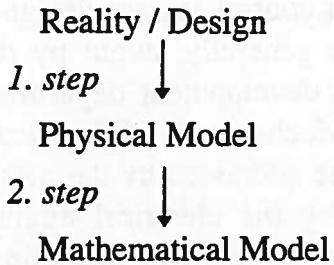
A major reason for this is the mathematical and theoretical emphasis given to these topics by research and education, which is aimed at improving the capabilities of the method itself (more depth), whereas only little effort has been put in increasing the industrial application of these methods (more width). This is also reflected by the amount of complex mathematics found in many textbooks. As a result of this, a huge gap exists between theory and daily engineering practice.

In the field of modal-testing, the focus on improvements in measurement accuracy can be understood from the fact that the driving forces are the aeronautic and automotive industries. In both these branches, the modal density (number of modes per frequency range) of the test object is extremely large, and the measurement set-up is very demanding. This explains why a lot of effort is put into improving the accuracy of the measurements; this includes data acquisition, digital signal-processing, and modal parameter-estimation. This situation is completely different from the situation in the field of equipment and consumer products, where often only a few modes are of major importance, and even an elementary modal-extraction technique suffices.

Papers that illustrate the use of modal analysis as a design tool to minimise unwanted dynamic effects in machines and products are very rare. The use of the modal properties of a high precision positioning device in the design of an optimised suspension system is described in [EIJ88]. The correspondence between controllability and observability (which are both important aspects in modern control theory), and the position of driving and response points of a vibration mode of a positioning device, are described in [SET89,SET92]. The authors use the knowledge about vibration modes to eliminate unwanted resonances that endanger the stability of the control loops.

Apart from the experimental technique, the numerical techniques have also evolved during the last decades and the development of computer technology has resulted in powerful software tools for the analysis of dynamic systems. However, one must be aware of the fact that these tools support only one part of the modelling process. In general, the modelling of certain phenomena that take place in a machine (for example heat distribution in a personal computer or the static deflection of a machine component due to a force) can be divided into two major steps [HEE72] :

¹Experimental determination of natural frequencies and mode-shapes of a structure.



In the first step, the real structure or design drawing of a structure or mechanism needs to be translated into a physical model, which is a simplification of the reality that contains all relations considered to be important to describe the phenomenon. Once this physical model has been derived, the second step consists of translating this physical model into a mathematical model, which can be solved either by hand or by means of a computer program. This second step, in general, is relatively straightforward, because it is based on existing approaches and rules. The first step, however, is much more demanding, and requires a lot of engineering judgement.

Computer programs can only assist during the second step of the modelling process (for example automatic FE mesh-generation based on a solid modelling description of the geometry), and new features are continuously being added by the suppliers of the analysis software. Obviously these suppliers focus entirely on their product, and thus on the second step of the modelling process. However, also in the area of research and education, most emphasis is put on the second step.

The questions :

- which analysis must be carried out
- how should the results be interpreted
- what sort of physical model gives a reasonable balance between accuracy and required effort (=time)

are only seldom addressed.

A very comprehensive and systematic approach for the creation of a physical model on the basis of which the major dynamic effects of cam mechanisms can be analysed is described in [KOS73]. The problem of optimising servo parameters of a servo system in which the end effector is compliantly linked to the servo motor is discussed by [BOU83] and [GRO91], using a very simple 2-DOF system.

A discrepancy between available theory and engineering practice can also be observed in the field of control theory. This can be concluded from the fact that despite new development in control theory, many high-precision manipulators such as CD-modules and wafer-steppers achieve sub-micron accuracy on the basis of traditional (SISO) PID controllers.

Apart from the gap between theory and practice in each of the two individual disciplines, the integration of machine dynamics and control system design is also limited. In educational institutes, the two topics are generally taught by different departments or even different faculties. In industry, development departments are mostly organised according to the disciplines "Mechanical", "Electrical" and "Software" engineering. Machine dynamics is an issue addressed by the mechanical engineer, whereas the control system is designed by the electrical engineer. In combination with the different terminology employed by the two disciplines (e.g. pole vs. resonance, zero vs. anti-resonance), the co-operation is often limited.

The lack of integral knowledge of machine dynamics, control and the interaction between these two topics is a serious threshold in finding the optimal solution to a mechatronic design problem.

1.3. Scope and Purpose

The major aim of this thesis is to preserve the knowledge in the field of machine dynamics, control design and modelling that has been gathered during the past ten years while working on various mechatronic projects ranging from compact disc modules to wafer-steppers and component-mounters. The insights into the effect of machine dynamics on the performance of servo-controlled positioning devices, and the working method aiming at "predictive modelling" of these effects, are considered useful for new colleagues and other mechatronics engineers.

This thesis aims at bridging the gap between existing theoretical knowledge in the field of machine dynamics and control, and the practical application of this knowledge during the design (or the experimental evaluation) of a product or machine. Both the designer and the test engineer, in most cases, are not experts in dynamics, and certainly not daily users of matrix calculus. Therefore an attempt is made to present the theoretical knowledge in a way that is believed to give more insight and feeling for dynamics in mechatronic systems, and especially for the effect of mechanical resonances on the stability of the control system.

The idea is to show that a basic understanding of machine dynamics suffices to interpret complex mechanical vibrations. Moreover, in combination with basic control theory it is possible to derive the typical patterns that can be observed in an open-loop frequency response of a mechanical servo-system including resonances, and to draw conclusions with respect to the effect of these resonances on the stability of the control loop. Based on the idea that the controlled system must satisfy certain disturbance rejection and bandwidth criteria, design guidelines can be given for the mechanical system such that the chance of realising the required bandwidth without introducing stability problems is maximised. By using a step-wise modelling approach it is possible to investigate and "predict" these phenomena during the

design phase, and to make design decisions which take the dynamics and control aspects into account. This thesis addresses the expert as far as the scientific background of the presented approach is concerned, and the user (non-specialist) as far as the design rules are concerned.

In this thesis the influence of geometrical non-linearities is not considered, because apart from multi-purpose robot configurations, the majority of devices such as compact disc drives, wafer-steppers, and component-mounters, is not dominated by this type of non-linearity. Furthermore, even in configurations that include geometrical non-linearities, a significant number of problems can be tackled on the basis of linearised descriptions. Other non-linear effects, such as backlash and friction, are also not considered: first, because most problems of machine dynamics in mechatronics systems are caused by mass and flexibility effects, which can be solved on the basis of a solid understanding of machine dynamics of linear systems; second, because apart from some rules of thumb for certain special situations, backlash and friction effects in general are badly understood, and a lot of work still needs to be done before design and test engineers can really profit from the theoretical insight. Finally, due to the unpredictability of these effects and the lack of general control strategies to cope with these non-linearities in mechatronics systems, most effort is normally focused on avoiding non-linearities by means of a proper mechanical design.

1.4. Preview

The basic questions that are addressed in this thesis are :

- What sort of dynamic effects are important in mechatronic devices ?
- How can the dynamics of a complex system be described and understood ?
- What is the influence of mechanical resonances on the stability of a control loop ?
- Which design rules can be given to minimise the destabilising effect of machine dynamics ?
- How can one predict the machine dynamics in an industrial way, such that simulation and modelling is experienced as an effective design tool ?

Chapter 2 discusses why machine dynamics is an important aspect in the performance of positioning devices, and how it manifests itself. Furthermore, a glimpse is given of three characteristic dynamic effects that can often be observed.

In the understanding of the dynamic properties of a machine the concept of modal description plays a central role. After a short resume of the dynamics of linear systems, Chapter 3 presents a new graphical representation of this concept, followed

by a number of illustrative applications that show the ease of its use in understanding often-observed phenomena. This graphical approach is believed to facilitate a better understanding of the modal concept, and to enable a very intuitive link between the modal and the physical domains.

Chapter 4 addresses the question how the desired motion of a manipulator and one single mode of vibration interact, and what effect that has on the frequency response function and servo stability. The basic shapes of a mechanical frequency response function that are found in this discussion are then linked to the three major dynamic effects, and design guidelines are derived.

Chapter 5 suggests a step-wise approach to tackle the problem of machine dynamics during the design process. This approach is based on the idea of assisting the design engineer in making the proper design decisions from a dynamics point of view. Directly linked to this idea is the modelling philosophy of doing the right sort of analysis (accuracy \leftrightarrow speed) at the right time in the design process, which can be achieved by a step-wise modelling approach. Practical modelling issues, including substructuring, modal truncation, and mixed mechanics/control simulations, are also addressed.

Chapter 6 describes a case study of a compact disc module; this is intended to illustrate the previous topics in general, and the modelling approach in particular.

Vibration of the machine frame as a result of the driving forces has proved to be a major issue in the settling time of servo devices. Despite its importance, this issue has been moved to Appendix A in order to enhance the structure of this thesis. In this appendix, guidelines are given to estimate the effect of frame vibrations on settling behaviour. Also, the difference between direct drive and indirect drive concepts is discussed with respect to this issue. The chapter ends with a systematic overview of approaches to minimise the deteriorating effect of these frame vibrations.

Finally, Chapter 7 states the conclusions of this work, with some concluding remarks.

2. Mechanical Servo Systems

This chapter is intended to familiarise the reader with the various dynamic aspects encountered in the design and analysis of a mechatronic positioning-system. The mechanical servo-systems considered in this thesis in general have the task to achieve a certain positional relation between two or more components of a system. This can be :

- fixed (relative) position
- point-to-point motion
- tracking/contouring

In the case of a video disc player (Fig.2.1) or a compact disc player, the servo system must ensure that the laser spot remains focused on the track of the disc (required accuracy typically $0.2\mu\text{m}$) although the track moves vertically due to unflatness of the disk, and horizontally due to eccentricity of the clamping hole (about $200\mu\text{m}$).

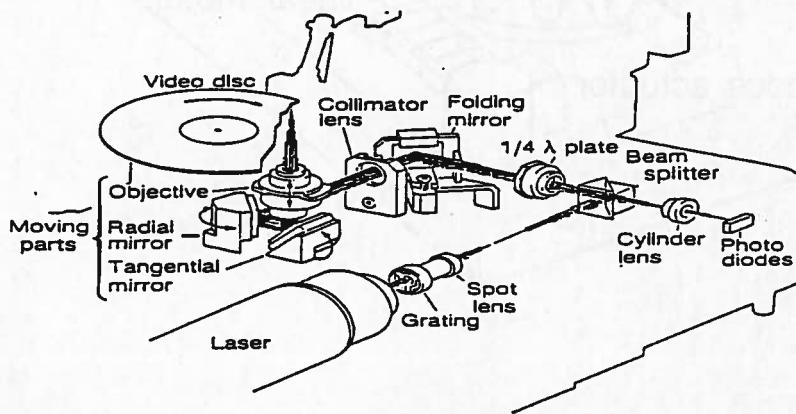


Figure 2.1 Video Disc module

Another interesting machine is the optical wafer-stepper, which is used in the photolithographic production of integrated circuits. The main task of this machine is the projection of an image (reticle) via an attenuation lens on to a Silicon wafer. This attenuation step is required in order to achieve the extreme requirements with respect to the critical dimensions of the fine-pitch pattern, which are in the sub-micron range. The field size of the projected image is about 10x10mm, and after exposure of

one such field, called die, the wafer is moved to the next location ("wafer-stepper"). In this way a matrix of identical patterns is generated on the wafer.

The "XY stage" that is used to manipulate the wafer under the lens is shown in Fig.2.2. The wafer is clamped to the vacuum chuck, which is supported by an air-bearing running on a huge granite block and is driven by an H-configuration of three linear motors. Interferometers are used to measure the position of the chuck with respect to the machine frame.

The throughput of such a machine is one of its main performance criteria, and depends very much on the acceleration and speed with which the XY stage moves from one position to the next. However, the throughput also depends very much on the level and settling time of the vibrations that are present after approaching the new exposure position. Obviously, an accurate exposure process is only achieved if the level of vibration of all relevant machine parts is sufficiently low.

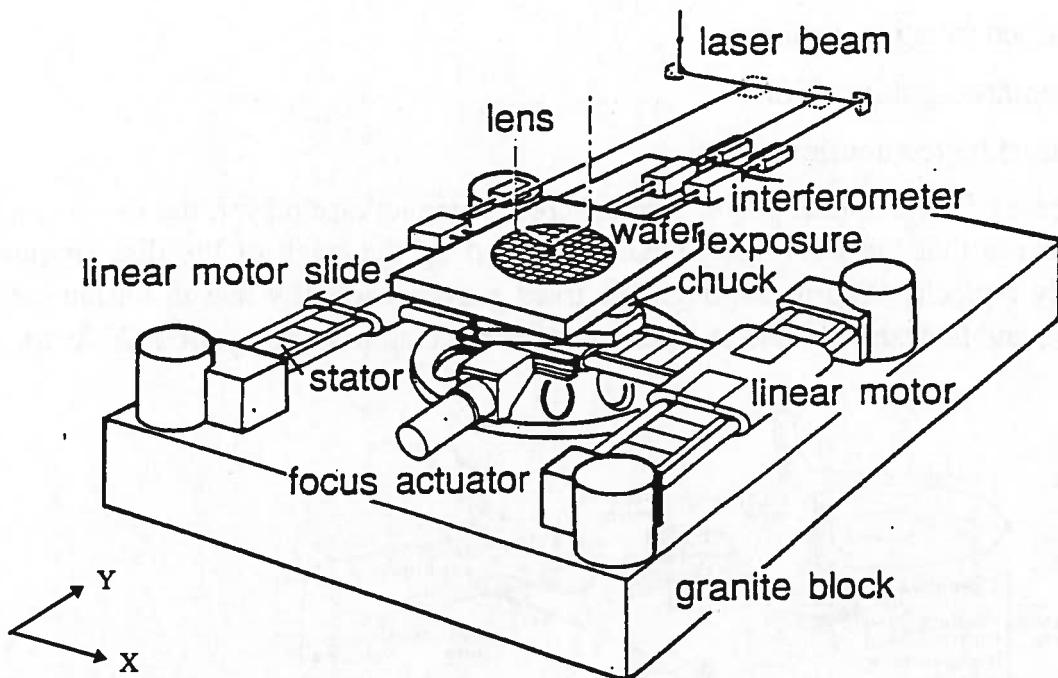


Figure 2.2 ASM-L wafer-stepper

In the third application area - tracking/contouring - the motion of the manipulator must "exactly" match the desired trajectory during the entire motion. NC machines and scanners typically fall into this category. Figure 2.3 shows an example of a machine which is used for the inspection of printed circuit boards. In this machine the printed circuit board is placed on a carrier, which during the inspection has to move at a constant speed of 20mm/s, while the position error typically has to remain smaller than 2 μ m.

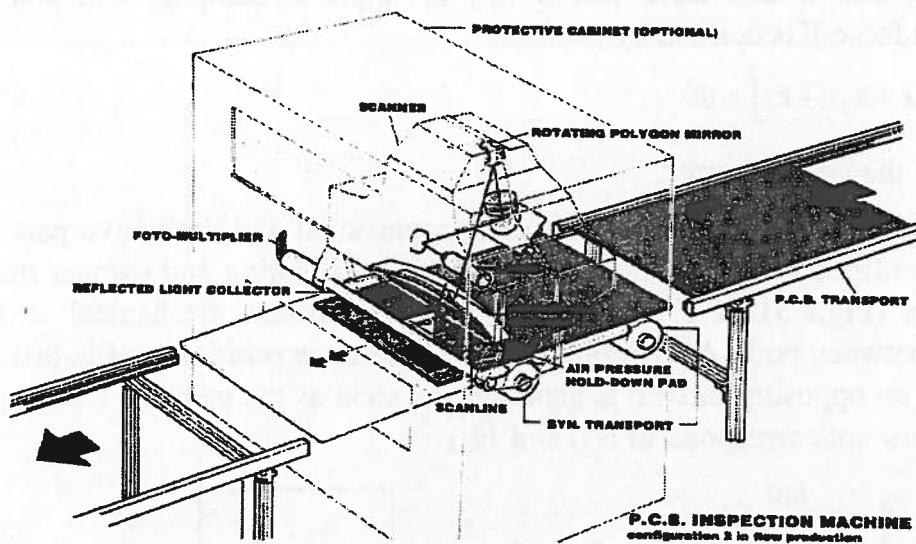


Figure 2.3 PCB Inspection Machine

2.1. Basic Control Aspects

A block diagram representation of a typical mechanical servo-system is shown in Fig.2.4. As mentioned, the main task of the system is to achieve a desired positional relation between two or more components of the system. Therefore, a sensor measures the position which is then compared to the desired value, and the resulting error is used to generate correcting forces. In most systems, the “actual output” (for example position of a robot end-effector) cannot be measured directly, and the feedback will therefore be based on a “measured output” (for example, an encoder signal at the motor). It is important to realise that these two outputs can differ, first due to resilience in the mechanical system, and second because of geometrical imperfections in the mechanical transmission between motor and end-effector, or due to imperfections of the encoder.

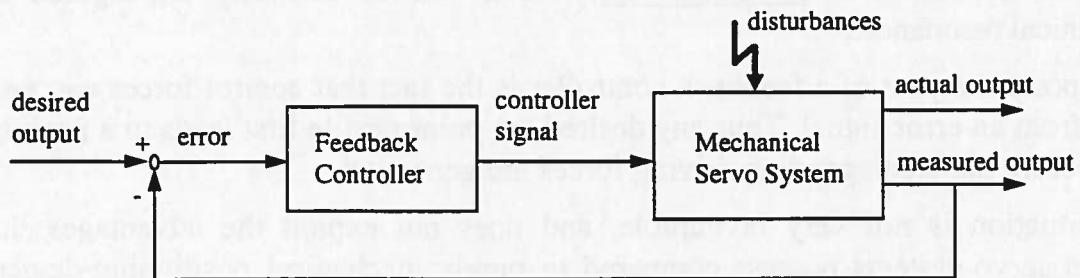


Figure 2.4 Basic Elements of Mechanical Servo System

The basic ingredient of most feedback control-systems is a PID controller, which is a combination of a proportional gain (P), an integral control (I) to enhance steady state

behaviour, and a derivative action (D) to improve damping and stability. The correction force F is defined by :

$$F = k_p \varepsilon + k_d \dot{\varepsilon} + k_i \int \varepsilon dt \quad (2.1)$$

where ε is the position error.

It is illustrative to see that basically the proportional and derivative part of such a position control loop is very similar to a mechanical spring and damper that connect two points (Fig.2.5). If c and d represent the constant mechanical stiffness and damping between point A and point B, and a reference position profile $h(t)$ is applied at A, then an opposing force F is generated as soon as the position x and speed \dot{x} of point B does not correspond to $h(t)$ and $\dot{h}(t)$.

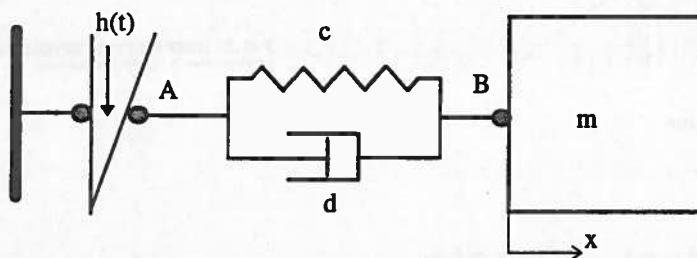


Figure 2.5 Basic Elastic Structure

The resulting spring/damper force is equal to :

$$F = c[h(t) - x] + d[\dot{h}(t) - \dot{x}] \quad (2.2)$$

Thus, in this example a PD position-control loop can be treated in the same way as a mechanical system with a spring ($c=k_p$) and damper ($d=k_d$)¹. The most important difference is the fact that a mechanical spring/damper is a passive element that builds up a force between its end-points A and B due to a change of distance or speed of these two points, whereas firstly a control loop is an active element, and secondly the servo forces that are applied between points A and B can be based on a measurement at a different location. These properties are very essential since they introduce the issue of servo stability, which can be seriously endangered by mechanical resonances.

An important aspect of a feedback controller is the fact that control forces can only result from an error signal. Thus any desired set-point profile first leads to a position error before the corresponding driving forces are generated.

This situation is not very favourable, and does not exploit the advantages that modern servo-systems possess compared to purely mechanical positioning-devices such as cam mechanisms. Nowadays, most modern industrial servo-systems have not only a feedback section, but also a feedforward section, as indicated in Fig.2.6.

¹ Due to this analogy, the servo parameters k_p and k_d are also called "servo stiffness" and "servo damping".

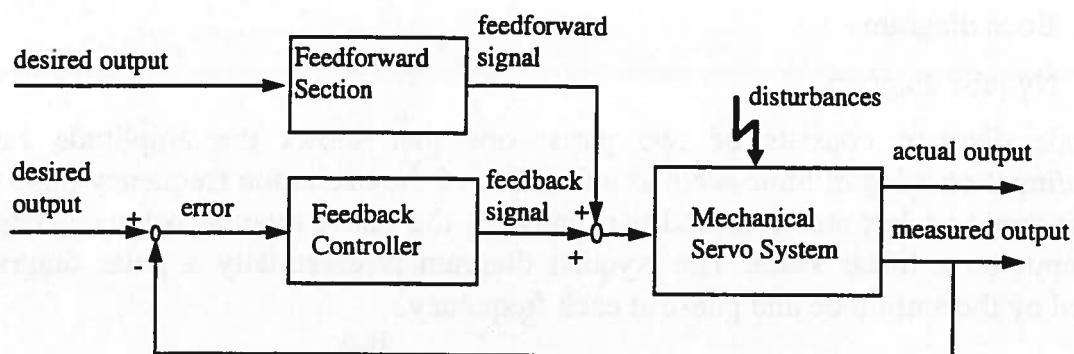


Figure 2.6 Mechanical servo system with feedback and feedforward control

In the feedforward section, control signals are derived from the desired output (position, speed, and acceleration) using a model of the mechanical servo-system. In many industrial controllers, such a model assumes that the servo system behaves as one rigid body on which coulomb and viscous forces are applied by the guiding system. The benefit of this approach is the fact that the nominal driving forces (acceleration forces plus forces to overcome coulomb and viscous friction) corresponding to a certain motion profile are applied directly to the system, and do not need to be generated via a position error.

In practice, a feedforward section gives a significant performance improvement in case of point-to-point and tracking applications, but a feedback section will always be necessary. First, because the model used in the calculation of the feedforward signal is not a perfect representation of the actual system, and second because of the presence of unknown disturbances. Furthermore, some systems (for example a magnetic bearing) need to be stabilised by a feedback controller.

Tuning the parameters of the feedforward section is generally quite straightforward, as in most cases it only involves an estimation of the total moving inertia of the system (multiplied by some system constants such as amplifier gain and motor constant), and experimental determination of coulomb and viscous friction effects of the system. Alternatively, the feedforward parameters can be tuned by relevant test motions during which the parameters are adjusted such that the monitored position error is minimised.

The design of the feedback control section is more complicated, for it brings up the issue of servo stability. Despite the emphasis in educational environments on methods such as root loci and optimal control, practising engineers generally accomplish the feedback design and analysis on the basis of the frequency response. One of the major benefits of this approach is the close link to experimental information that can be obtained by exciting the system with sinusoidal inputs of varying frequency and measuring the amplitude and phase of the output. Frequency response data can be plotted in various formats, two of which will be shortly introduced here, because they are used later.

- Bode diagram
- Nyquist diagram

A Bode diagram consists of two parts: one that shows the amplitude ratio output/input on a logarithmic scale as a function of the excitation frequency (also on a logarithmic scale); and a second part showing the phase relation between output and input on a linear scale. The Nyquist diagram is essentially a polar diagram defined by the amplitude and phase at each frequency.

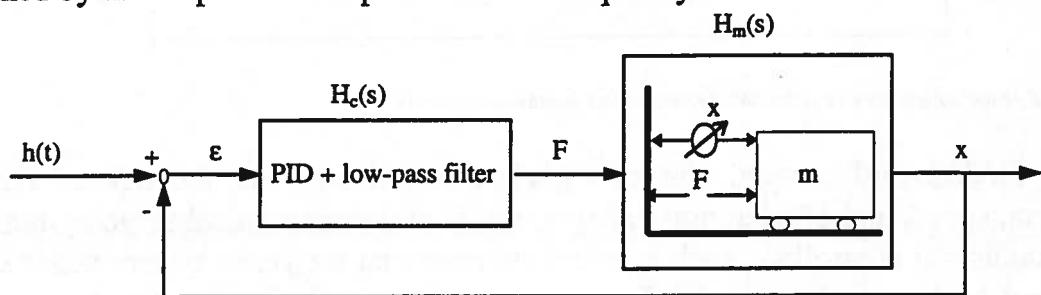


Figure 2.7 Mechanical servo system

As an illustration one can consider the very basic servo system (Fig.2.7) consisting of a rigid mass m that is driven by the servo force F . The position of the mass is measured relative to the fixed world. A PID controller in series with a second order low-pass filter operates on the position error ϵ and generates the servo force F . The Bode and Nyquist diagrams of the open loop transfer function $H_c * H_m$ of this system are shown in Fig.2.8.

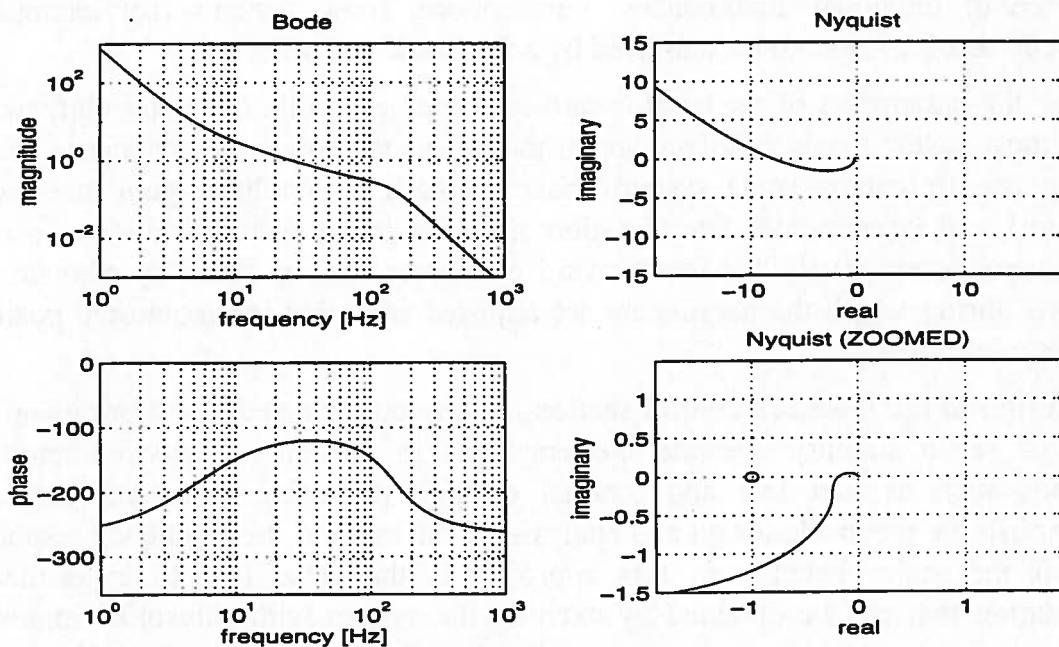


Figure 2.8 Bode and Nyquist diagrams of the open loop frequency response

The feedback controller must satisfy certain stability criteria, a complete and mathematically correct treatment of which can be found in most textbooks on control theory, for example [FRA88], [DOF86]. An engineering approach to stability evaluation is the so-called “Left Hand Rule” [COO79], which reads :

If a system contains only stable elements in the open loop, then the closed loop system will be stable if the point (-1,0) in the Nyquist diagram lies on the left hand side of the open loop response when it is run through in the direction of increasing frequency.

In order to quantify the level of stability, two criteria have been introduced :

- gain margin
- phase margin

which measure how close the open-loop response approaches the point (-1,0) in the Nyquist diagram.

Assuming that at a certain frequency the open-loop phase equals -180°, then the gain margin is the factor by which the amplitude ratio at that frequency differs from “1”. Complementarily, the phase margin is the amount by which the open-loop phase exceeds -180° at the frequency for which the amplitude ratio equals “1”.

If the “left hand” stability criterion is applied to Fig.2.8, one can conclude that the system is stable. The gain margin equals a factor 1/8 (at 5Hz), and a factor 10 (at 150Hz), whereas the phase margin amounts to about 60° (at 30Hz).

2.2. Specifications

Specification of a feedback controller is very closely linked to disturbance rejection, especially in modern controllers that incorporate a feedforward section. As mentioned before, if the feedforward control signal is based on an exact representation of the actual (stable) positioning device and all initial conditions are zero, the nominal set-point motion can be achieved without any feedback action. Therefore, the required performance of the feedback section, which is generally expressed in terms of bandwidth², depends very much on the disturbances that act on the system.

These disturbances can be very different, and vary from application to application:

²The definition of bandwidth as the first 0dB crossing of the open-loop frequency response is gaining popularity, as under the assumption of a standard PID controller it gives a direct indication of the low-frequency disturbance rejection even in the presence of resonance phenomena. The traditional definitions of bandwidth as the frequency at which the closed-loop magnitude plot equals -3dB, or the phase plot equals -90°, can sometimes be over-optimistic when judging the low-frequency disturbance rejection on the basis of that bandwidth, because this definition is very sensitive to the presence of resonances.

- In the case of a Compact Disc Player the eccentricity of the disc itself with respect to the shaft of the disc motor leads to a sinusoidal radial motion of the track (frequency about 10 Hz) that is being read. The amplitude of this radial motion may be as large as 0.2mm, whereas the tolerable radial position error may not exceed about 0.2 μ m. The servo system must be capable of following that sinusoidal motion while keeping the position error smaller than 1/1000 of the amplitude of that motion.
- Load fluctuations due to cogging forces of the motor.
- Variable coulomb and viscous friction forces or other imperfections of the guiding system.
- Random floor vibration in the case of very sensitive equipment, such as wafer-steppers or wafer-inspection machines. In the current and next generation ICs, line-widths are reduced to about 200nm, which implies that the various error budgets (vibrations being only one of the many issues that contribute to the overall error) are very tight. When talking about allowable vibration levels of a few nm it is not amazing that dynamics becomes very significant.
- Harmonic excitation forces due to the presence of pumps or ventilators.
- Shocks and bumps (for example portable Compact Disc Players that are used during jogging).
- Acoustic excitation, which gains more and more attention because it is very difficult to protect a machine from this kind of disturbance which, furthermore, contains many high frequency components that normally cannot be suppressed by the control system.

2.3. Interaction Dynamics and Control

Basically, machine dynamics can have two deterioration effects in mechanical servo systems:

- Mechanical resonances can endanger servo stability, and thus limit the bandwidth and the amount of disturbance rejection.
- Vibrations can lead to positional errors at the end of a set-point motion, or during a tracking motion.

To illustrate this, the measured frequency and time response of an actual machine will be discussed. The control consists of a PID feedback section and a feedforward section. Figure 2.9 shows the Bode and Nyquist diagrams of the open-loop frequency response. Clearly one can see that, due to the mechanical resonances, the overall

gain and thus the bandwidth and disturbance rejection cannot be increased without endangering the stability.

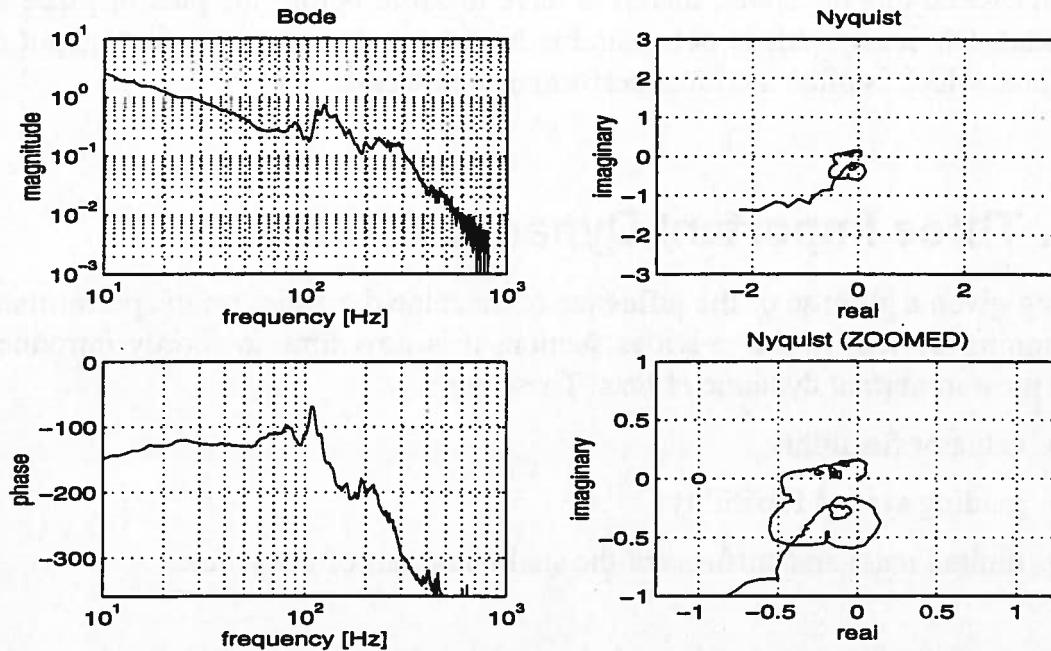


Figure 2.9 Measured open-loop frequency response (Bode and Nyquist diagrams)

The time response of the system to a third-order set-point profile is shown in Fig.2.10. In the absence of any machine dynamics, and assuming a well-tuned feedforward section, the output would have followed the input exactly without any position error. Due to the dynamics, however, one can observe vibrations during the set-point motion and even after the end of the motion profile.

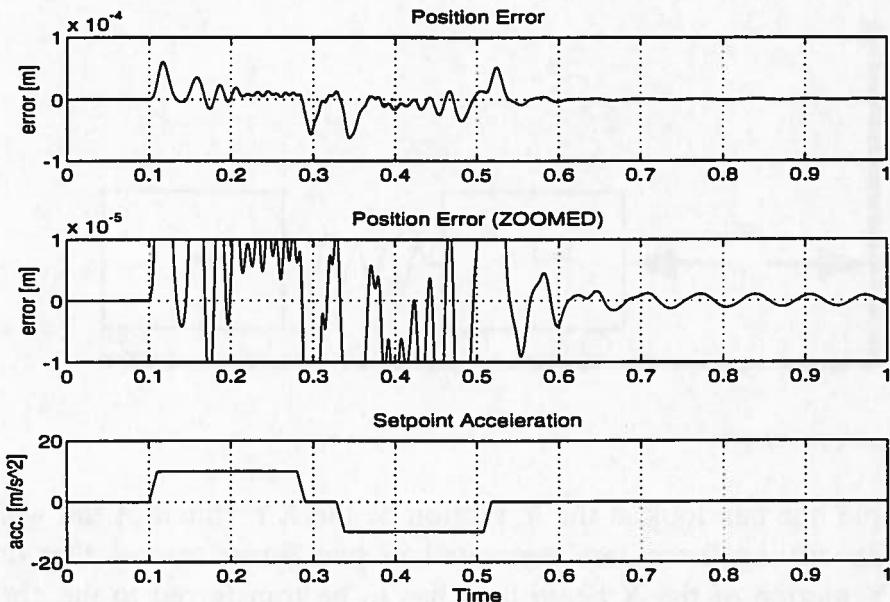


Figure 2.10 Measured time response due to a cubic set-point profile

In pick-and-place applications a component can only be picked and placed correctly if the position error is below a certain threshold. Vibrations at the end of a set-point, which exceed this threshold, therefore have to settle before the pick or place action can start. Obviously, this is not desirable because it decreases the throughput of the machine, which is often a critical performance criterion.

2.4. Three Important Dynamic Effects

Having given a glimpse of the influence of machine dynamics on the performance of positioning devices in the previous section, it is now time to shortly introduce the three most important dynamic effects. These are :

- actuator flexibility
- guiding system flexibility
- limited mass and stiffness of the stationary part of a machine

2.4.1. Actuator Flexibility

In a servo system a force or torque is applied in order to achieve a certain displacement or rotation of the driven system, which can be a single component or a complex driving system including couplings and transmission. The basic characteristics of what is called “actuator flexibility” is the fact that in the frequency range of interest (typically $0-10^*$ bandwidth) the driven system no longer behaves as one rigid body (Fig.2.11) due to compliance between the motor and the load.

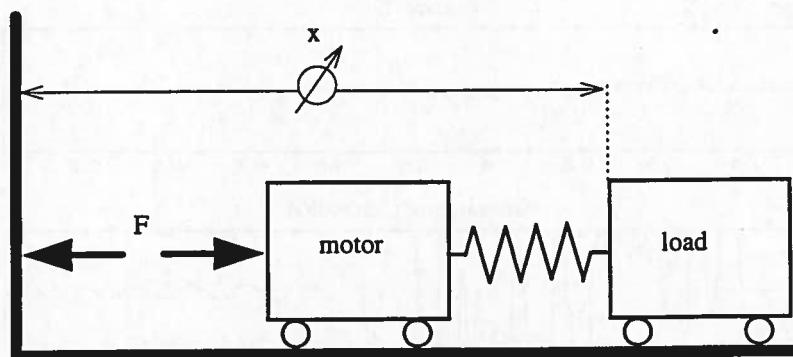


Figure 2.11 Actuator Flexibility

As an example one can look at the Y motion of the XY chuck of the wafer-stepper (Fig.2.2). The driving forces are generated in two linear motors that drive the X beam. The Y motion of the X beam then has to be transferred to the chuck, which holds the wafer. Because the bending stiffness of the X beam and the guiding

stiffness between X beam and chuck is limited, the Y servo will exhibit the dynamic behaviour that is characterised by "actuator flexibility".

2.4.2. Guiding System Flexibility

The second category of dynamic phenomena results from the limited stiffness of the guiding system in combination with the fact that the device is driven in such a way that it has to rely on the guiding system to suppress motion in an undesired direction (in case of a linear direct drive system this occurs if the driving force is not applied at the centre of gravity).

In general, a rigid actuator possesses six degrees of freedom, five of which need to be suppressed by the guiding system in order to leave one mobile degree of freedom. In the present discussion a planar actuator with three degrees of freedom will be considered (Fig.2.12).

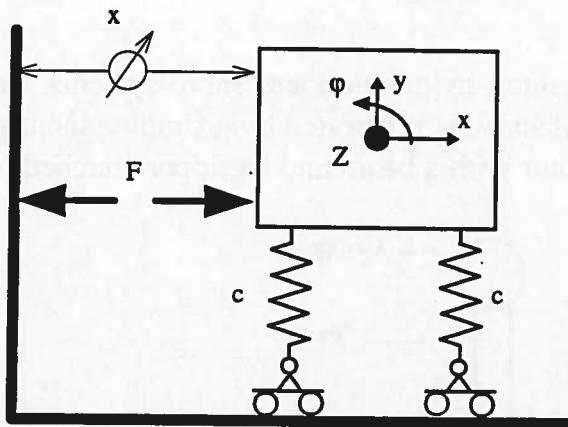


Figure 2.12 Planar actuator with guiding system flexibility

The carriage is free to move in the guiding direction (x), whereas the perpendicular displacement (y) and the rotation (φ) is prevented via two fixtures with limited stiffness c . The limited support stiffness and the inertia properties of the actuator will result in two resonances, which can be characterised as perpendicular mode and rocking mode³.

Every actuator has some sort of guiding system in order to suppress certain DOF, and thus possesses guiding modes. However, whether this leads to dynamic problems depends very much on the location of the driving force (and the sensor). By choosing the proper location of the driving force (centre of gravity) one can avoid excitation of these modes, whereas the location of the sensor influences the

³ When the centre of gravity Z is located symmetrically between the two identical supports, the DOF y and φ are completely decoupled, thus resulting in a pure y mode and a pure rotational mode about the COG. If this condition is not met, both modes would involve the DOF y and φ .

effect of such a mode on the servo stability where excitation of the mode could not be avoided.

In general it should be attempted to design the actuator (mass distribution and location of driving force) such that it will perform the desired motion even in the absence of the guiding system.

This phenomenon can be illustrated by looking at the X motion of the XY stage of the wafer-stepper (Fig.2.2). The location of the wafer must be known with a very high accuracy, so the position of the chuck is measured as close as possible to the wafer in order to avoid Abbe-errors that result from imperfections of the guiding system. Due to design limitations it was not possible to apply the driving force at the centre of gravity but it had to be applied at a certain distance below the centre of gravity. Consequently, the situation shown schematically in Fig.2.12 has been created, and as a result of this the X servo suffers from this phenomenon. Fortunately, the distance between the driving force and the centre of gravity is limited, and therefore the deteriorating effect on the servo stability could be minimised.

The described effect is not limited to translational servo-systems, but can also be observed in rotational servo-systems, as illustrated by a simple planar pick-and-place manipulator consisting of a motor with a beam and a gripper attached to it (Fig.2.13).

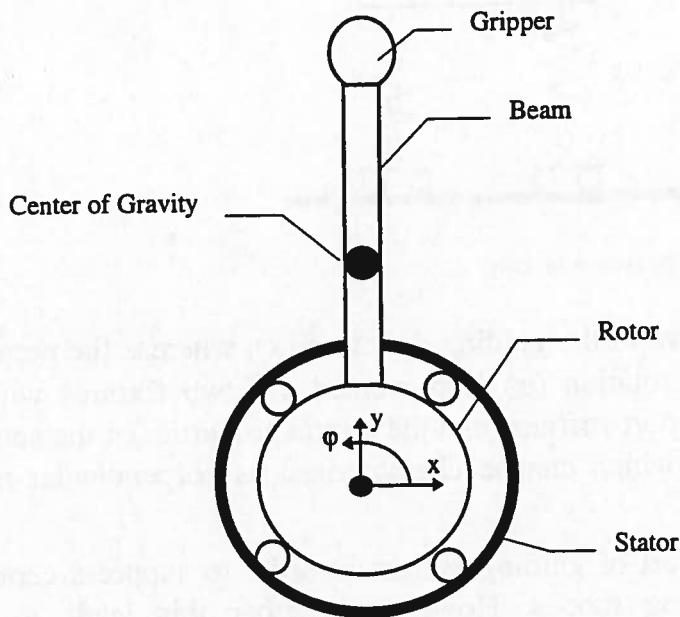


Figure 2.13 Planar rotational pick-and-place manipulator

Neglecting the flexibility of the beam (= actuator flexibility), the manipulator possesses two parasitic modes due to the limited stiffness of the bearing between rotor and stator in combination with the inertia of the rotor, the beam, and the

gripper. Assuming a symmetric configuration, one of these modes is a pure y mode, whereas the other mode involves both the x and φ DOF. When driving the system the inertia forces will result in reaction forces on the bearing (centre of gravity does not coincide with the centre of the bearing), and will excite the $x\text{-}\varphi$ mode. This can also be understood from the fact that in the absence of the bearing, application of a torque would lead to a rotation about the centre of gravity and not the centre of the bearing. This is not the desired motion, so it is suppressed by the guiding system at the risk of introducing dynamic problems.

2.4.3. Limited Mass and Stiffness of Stationary Machine Part

The last category of dynamic phenomena results from the limited mass and stiffness of the stationary part of a mechanical servo-system. In contrast to many textbooks on mechanics and machine dynamics, but in analogy with the electric discipline, it is good practice always to look at the combination of driving force on the moving part, and reaction force on the stationary part, of a positioning device. When doing so, one has to consider what the effect of the reaction force on the systems performance will be. In the discussion of the previous two dynamic phenomena, the stationary part of the machine was assumed to be infinitely stiff and heavy, and therefore the effect of the reaction force was negligible. However, in general the stationary part is neither infinitely heavy, nor is it connected to its environment with an infinite stiffness, so the stationary part will exhibit a resonance that is excited by the reaction forces (Fig.2.14).

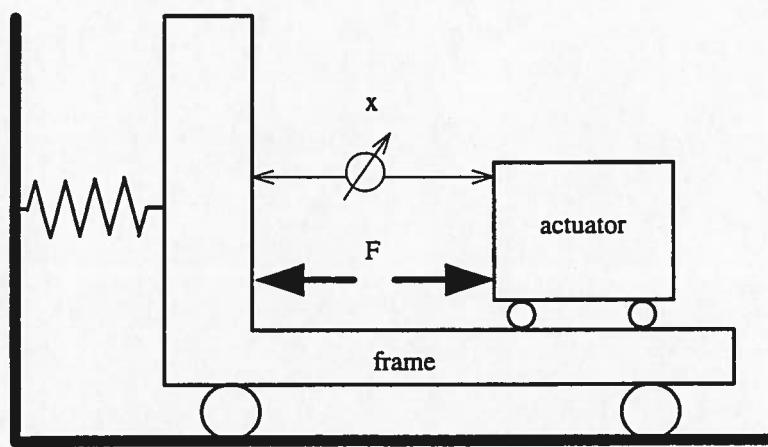


Figure 2.14 Limited Mass and Stiffness of Stationary Machine Part

Practice has taught that in a well-designed servo system the effect of such a resonance on the stability of the servo system is generally small; even then it can have a significant impact on the set-point response. More seriously, in a wide range

of applications, the time-domain performance is seriously limited by the effect of the vibrating frame that is excited by the servo system itself. Briefly, the phenomenon can be described as follows: the servo forces are applied on the actuator in order to generate a desired motion; however, the same forces also act on the frame and cause it to vibrate. At the end of the set-point the frame motion has not ended, and the actuator has to "follow" the motion of the frame. The acceleration force required to do so acts as a disturbance on the servo system, and inevitably introduces a position error. The effect of frame vibrations is even worse where the quality of positioning of the servo system is not determined by the position of the actuator relative to the frame, but by the position of the actuator relative to the world (for example a robot that has to pick a component from a pallet that is placed on the floor).

3. Modal Decomposition

To understand and describe the behaviour of a mechanical system in a quantitative way, one usually sets up a model of the system. The mathematical description of such a model with a finite number of DOF consists of a set of ordinary differential equations. Although in the case of simple systems, such as illustrated in Fig.3.1, these equations may be very understandable, in the case of complex systems the set of differential equations itself gives only very limited insight, and mainly serves as a basis for numerical simulations.

A very powerful tool, both numerically and experimentally, in understanding the dynamic properties of a mechanical system, sub-system or component, is the concept of "modal analysis", which will be discussed in this chapter. The theoretical foundation of modal analysis [EWI84] will be treated first. Then, the mathematics of one single decoupled "modal" equation will be translated into a graphical representation that includes all relevant data. This graphical approach is believed to facilitate a better understanding of the modal concept, and to enable a very intuitive link between the modal and the physical domain without the complications of the mathematical formalism.

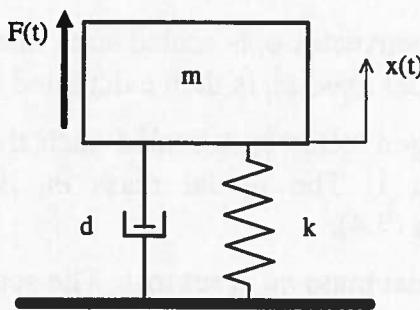


Figure 3.1 Elementary Dynamic System

3.1. Mathematics of Modal Decomposition

The general equation of motion of a linear mechanical system with a finite number of DOF, and without damping is :

$$M\ddot{x}(t) + Kx(t) = f(t) \quad (3.1)$$

in which \mathbf{M} and \mathbf{K} stand for the symmetric (semi-) positive definite mass and stiffness matrix, $\mathbf{x}(t)$ and $\ddot{\mathbf{x}}(t)$ represent the displacement and acceleration vectors, and $\mathbf{f}(t)$ denotes the vector of forces.

Generally this system of equations is coupled but it can always be decoupled by using a transformation based on the non-trivial solutions (the eigenvectors) of the following eigenvalue problem :

$$(\mathbf{K} + \omega_i^2 \mathbf{M}) \phi_i = \mathbf{0} \quad (3.2)$$

Solving the eigenvalue problem gives the eigenvalues $\omega_1^2, \omega_2^2, \dots, \omega_n^2$ and the corresponding eigenvectors or mode-shape vectors $\phi_1, \phi_2, \dots, \phi_n$.

These eigenvectors have the following orthogonality properties, or can always be chosen such that :

$$\phi_i^T \mathbf{M} \phi_j = 0 \quad (i \neq j) \quad (3.3)$$

For $i=j$ the result of the multiplication according to equation (3.3) yields a non-zero result, which is normally indicated as modal mass m_i^1 :

$$\phi_i^T \mathbf{M} \phi_i = m_i \quad (3.4)$$

Because only the direction but not the length of an eigenvector is defined, several scaling methods are used, all based on equation (3.4). Basically there are two options, one can either specify a certain scaling for the mode-shape vector and calculate the resulting modal mass m_i from equation (3.4), or one can choose a value for the modal mass m_i and scale the mode-shape vector ϕ_i such that equation (3.4) is satisfied.

The following three scaling methods are often used :

- $|\phi_i| = 1$ Each eigenvector ϕ_i is scaled such that its length is equal to 1. The modal mass m_i is then calculated from equation (3.4).
- $\max(\phi_i) = 1$ Each eigenvector ϕ_i is scaled such that its largest element is equal to 1. The modal mass m_i is then calculated from equation (3.4).
- $m_i = 1$ The modal mass m_i is set to 1. The scaling of the mode vector ϕ_i follows from equation (3.4).

The orthogonality properties also apply to the stiffness matrix \mathbf{K} . From (3.2), (3.3) and (3.4) one can find that :

$$\phi_i^T \mathbf{K} \phi_j = 0 \quad (i \neq j) \quad (3.5)$$

$$\phi_i^T \mathbf{K} \phi_i = \omega_i^2 m_i = k_i \quad (3.6)$$

¹ Italics are used to indicate the modal mass m_i and the modal stiffness k_i , whereas a normal font is used to identify a physical mass m , or a physical stiffness k .

Because the n eigenvectors ϕ_i form a base in the n -dimensional space, any displacement vector $\mathbf{x}(t)$ can be written as a linear combination of the eigenvectors. Let $q_i(t)$ be the response of the decoupled mode i , then the resulting displacement vector $\mathbf{x}(t)$ will be:

$$\mathbf{x}(t) = q_1(t)\phi_1 + q_2(t)\phi_2 + \dots + q_n(t)\phi_n \quad (3.7)$$

For one individual physical DOF x_k equation (3.7) is reduced to :

$$x_k(t) = q_1(t)\phi_{1k} + q_2(t)\phi_{2k} + \dots + q_n(t)\phi_{nk} \quad (3.8)$$

with ϕ_{ik} being the element of the mode-shape vector ϕ_i that corresponds to the physical DOF x_k .

The physical interpretation of equations (3.7), (3.8) is that any motion of the system at any time can be regarded as a combination of contributions of the various modes.

One can combine the eigenvectors in a matrix Φ and the co-efficients $q_1(t), q_2(t) \dots q_n(t)$ in a vector $\mathbf{q}(t)$, which leads to a more compact notation of equation (3.7) :

$$\mathbf{x}(t) = \Phi\mathbf{q}(t) \quad (3.9)$$

$$\text{with } \Phi = [\phi_1 \ \phi_2 \ \dots \ \phi_n] \quad \text{and} \quad \mathbf{q}(t) = \begin{vmatrix} q_1(t) \\ q_2(t) \\ \vdots \\ q_n(t) \end{vmatrix}$$

Substitution of $\mathbf{x}(t)=\Phi\mathbf{q}(t)$ in the original equation of motion (3.1) and premultiplication with Φ^T results in :

$$\Phi^T \mathbf{M} \Phi \ddot{\mathbf{q}}(t) + \Phi^T \mathbf{K} \Phi \mathbf{q}(t) = \Phi^T \mathbf{f}(t) \quad (3.10)$$

Using equations (3.3), (3.4), (3.5), and (3.6) finally leads to a set of uncoupled equations of motion that describe the contribution of each mode :

$$\begin{bmatrix} m_1 & & & & \ddot{q}_1 & & k_1 & & q_1 & & \phi_1^T \mathbf{f} \\ m_2 & & & & \ddot{q}_2 & & k_2 & & q_2 & & \phi_2^T \mathbf{f} \\ \vdots & & & & \vdots & & \ddots & & \vdots & & \vdots \\ m_i & & & & \ddot{q}_i & & k_i & & q_i & & \phi_i^T \mathbf{f} \\ \vdots & & & & \vdots & & \ddots & & \vdots & & \vdots \\ m_n & & & & \ddot{q}_n & & k_n & & q_n & & \phi_n^T \mathbf{f} \end{bmatrix} = \begin{bmatrix} \ddot{q}_1 \\ \ddot{q}_2 \\ \vdots \\ \ddot{q}_i \\ \vdots \\ \ddot{q}_n \end{bmatrix} + \begin{bmatrix} k_1 & & & & & & \\ & k_2 & & & & & \\ & & \ddots & & & & \\ & & & k_i & & & \\ & & & & \ddots & & \\ & & & & & k_n & \end{bmatrix} \begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_i \\ \vdots \\ q_n \end{bmatrix} \quad (3.11)$$

with $m_i = \phi_i^T \mathbf{M} \phi_i$ and $k_i = \phi_i^T \mathbf{K} \phi_i$

For the i -th modal co-ordinate q_i the equation of motion is :

$$m_i \ddot{q}_i(t) + k_i q_i(t) = \phi_i^T \mathbf{f}(t) \quad (3.12)$$

which is a simple second order differential equation similar to that of a single-mass spring system. Using basic formulae that are derived for a simple single-mass spring

system, one is now able to analyse the time and frequency response of all individual modes. Having done that, the total motion of the system can simply be obtained by summing the contributions of all modes.

Characterisation of the dynamics of a mechanical system in terms of frequency response behaviour plays a major role in the stability analysis of the control loop of a mechatronic device. In such an analysis one is typically interested in the transfer function (x_i/f_k), between a (measured) displacement x_i and a (servo)force f_k , which acts at the physical DOF x_k . Applying the principle of modal decomposition, any transfer function can be derived by first calculating the behaviour of the individual modes, and then summing all modal contributions.

The contribution of one single mode i to the transfer function (x_i/f_k) can be derived by first considering the response of the modal DOF q_i to a force vector \mathbf{f} with only one non-zero component f_k . In that case equation (3.12) is reduced to :

$$m_i \ddot{q}_i(t) + k_i q_i(t) = \phi_{ik} f_k(t) \quad (3.13)$$

After a Laplace transformation and some rearrangement, equation (3.13) yields² :

$$q_i(s) = f_k(s) \frac{\phi_{ik}}{m_i s^2 + k_i} \quad (3.14)$$

Once the modal response q_i is known, the response of the physical DOF x_i is found by a simple premultiplication with ϕ_{ii} (see equation (3.8)), which finally leads to the following expression for the contribution of mode i to the transfer function (x_i/f_k) :

$$\left(\frac{x_i}{f_k} \right)_i = \frac{\phi_{ik} \phi_{ii}}{m_i s^2 + k_i} \quad (3.15)$$

The overall transfer function (x_i/f_k) can be found by summation of the individual modal contributions, which all have the same structure :

$$\left(\frac{x_i}{f_k} \right) = \sum_{i=1}^n \left(\frac{x_i}{f_k} \right)_i = \sum_{i=1}^n \frac{\phi_{ik} \phi_{ii}}{m_i s^2 + k_i} \quad (3.16)$$

It is interesting to note that, whereas most control engineers usually think in terms of transfer functions that are placed in series and which consequently are multiplied with each other, modal decomposition results in a block diagram description of the mechanics with a parallel structure, as shown in Fig.3.2. In Chapter 4, which deals with modes and stability, the interference between modal contributions to a frequency response will be discussed in detail.

² In order to enhance readability, the Laplace-transformed quantities are not indicated by capitals.

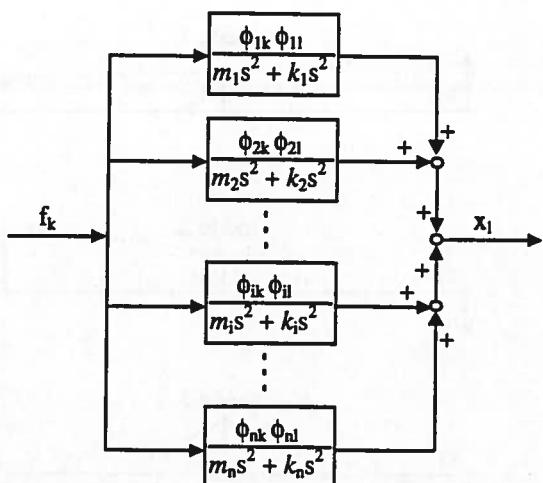


Figure 3.2 Block diagram representation of transfer function (x_i/f_k) with modal decomposition

3.2. Graphical Representation

Due to the equivalence with the differential equation of a single-mass spring system, equation (3.12) is often represented by a single-mass spring system on which a force $\mathbf{f}' = \phi_i^T \mathbf{f}$ acts, which is sometimes referred to as "modal force" [CRA81]. However, this representation implies an important loss of information because it neglects all information about the mode-shape vector. In this section a graphical representation of modal data [EIJ88,RAN88/3,RAN88/4] will be shown, which is very useful in understanding the concept of modal decomposition and modal superposition.

This graphical representation of the decoupled "modal" equation enables the user to apply the concept of "modal description" in a more intuitive way. Items such as the effective modal mass at a driving point, sensitivity analysis, or the application of a compensating force in order not to excite a mode, can easily be explained.

3.2.1. Lumped Mass Model

First, the graphical approach will be illustrated using an example of a lumped-mass model, consisting of three masses with intermediate flexibility (Fig.3.3a). For this system, the three mode-shapes are depicted in the traditional graphical representation (Fig.3.3b) which is found in many textbooks. In this representation, the physical DOF are located at fixed positions and the mode-shape displacement is indicated by the length and direction of an arrow.

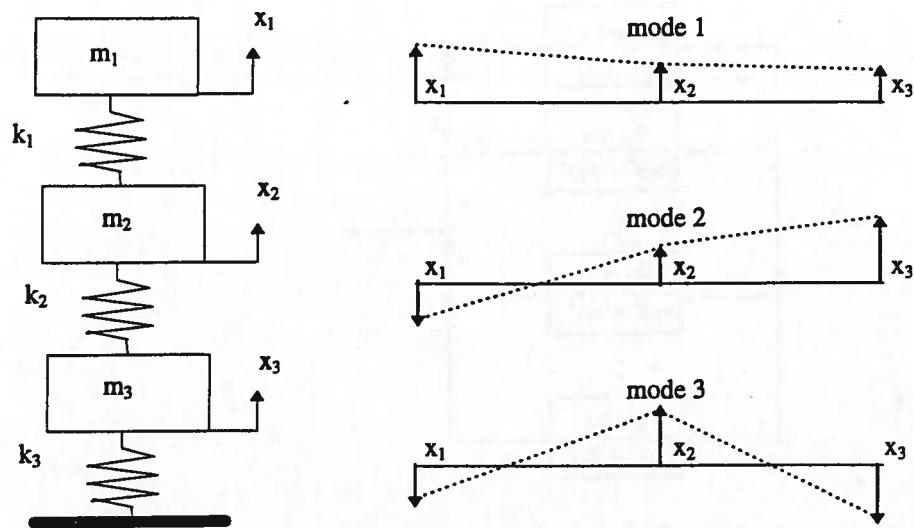


Figure 3.3 a. System b. Traditional graphical representation of modes

Alternatively, considering that for each mode the mode-shape vector defines a constant relation or transmission ratio between the various physical DOF, one could also represent a mode-shape by a lever (Fig.3.4). Per individual mode i , each physical DOF x_k is indicated on the lever at a position with respect to the point of rotation that corresponds to the amplitude and sign of that DOF in the mode-shape vector ($= \phi_{ik}$). Systems with no, very little, or proportional, damping exhibit real mode-shape vectors, and thus the various DOF reach their maximum values at the same moment of the cycle. Consequently, the respective DOF can only be in phase or in opposite phase. All DOF on the same side of the rotation point have identical phases, whereas DOF on opposite sides have opposite phases.

The modal DOF q_i can either be interpreted as the angle of the modal lever (Fig.3.4a) or as a displacement at a distance “1” from the pivot point (Fig.3.4b). In both cases the correct relation between the physical and modal domain exists. According to equation (3.8), any physical DOF x_k can be expressed in terms of the modal DOF q_i , as :

$$x_k(t) = \sum_{i=1}^n q_i(t) \phi_{ik} = \sum_{i=1}^n x_{ki}(t) \quad (3.17)$$

Thus the contribution of one single modal DOF q_i to the physical DOF x_k is represented by :

$$x_{ki}(t) = q_i(t) \phi_{ik} \quad (3.18)$$

In view of the more convenient graphical representation of modal mass and stiffness, the representation shown in (Fig.3.4b) will be used in all following chapters.

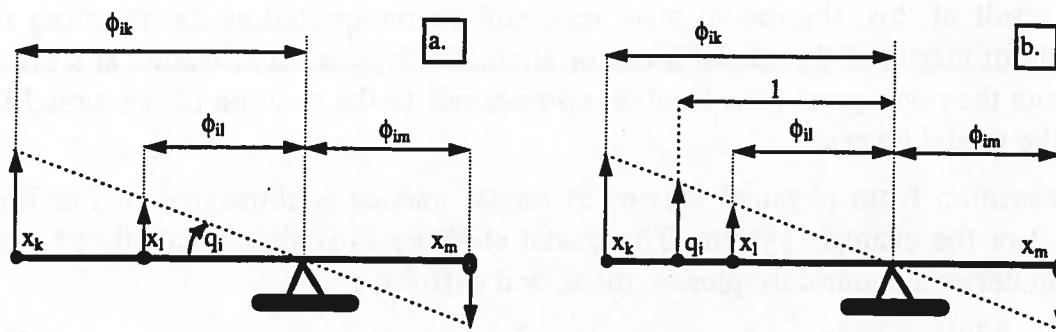


Figure 3.4 New graphical representation of mode i (a. q_i as angle of lever, b. q_i as displacement at distance "1")

Application of this new approach to the example system leads to the graphical representation shown in Fig.3.5.

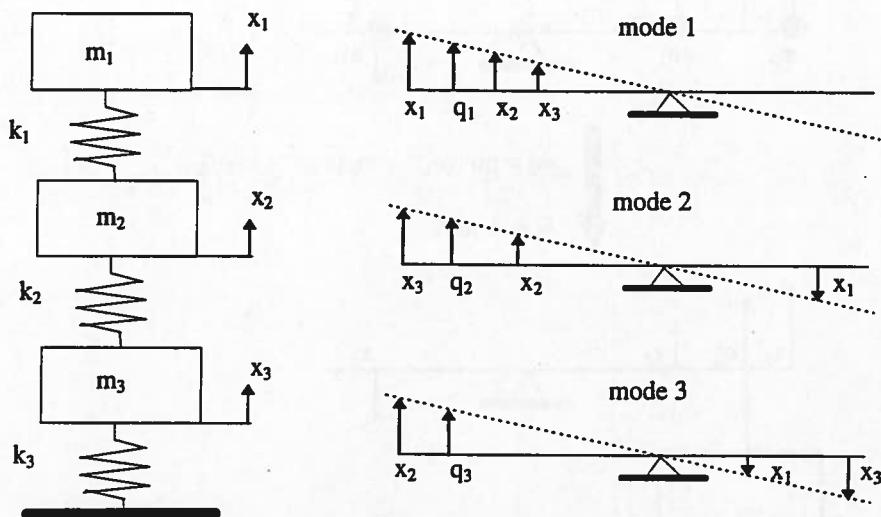


Figure 3.5 a. System b. New graphical representation of mode-shape

In the case of a lumped mass model, as in the previous example, it is possible to indicate at each physical DOF on the modal lever the corresponding physical mass, as shown in Fig.3.6a. The resulting mass moment of inertia J_i of the i -th modal lever then is:

$$J_i = \sum_{k=1}^n m_k \phi_{ik}^2 \quad (3.19)$$

This result is identical to the modal mass m_i found on the basis of equation (3.4), because the mass matrix \mathbf{M} is a diagonal matrix of physical masses m_k , and consequently the expression for the modal mass m_i yields :

$$m_i = \phi_i^T \mathbf{M} \phi_i = \sum_{k=1}^n m_k \phi_{ik}^2 \quad (3.20)$$

As a result of this, the modal mass m_i could be interpreted as the resulting mass moment of inertia of the modal lever, or alternatively as a mass located at a distance “1” from the pivot point (this location corresponds to the position of the modal DOF q_i on the modal lever).

The transition from physical masses to modal masses is illustrated in Fig.3.6 for mode 2 of the example system. The modal stiffness k_2 is simply calculated via the relation between natural frequency, mass, and stiffness :

$$k_i = \omega_i^2 m_i \quad (3.21)$$

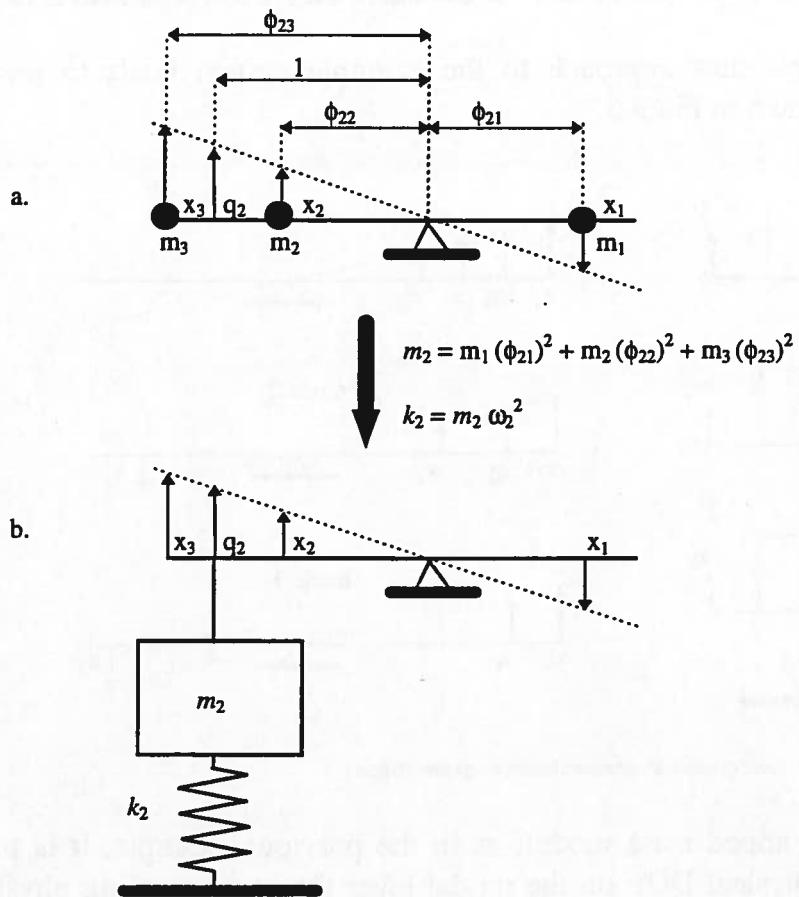


Figure 3.6 Graphical representation of mode 2 with (a.) lumped masses and (b.) modal mass and stiffness

3.2.2. General System

As a starting point, the concept of the “modal lever” representation has been introduced in the previous section for a lumped mass model. However, this is not a premise for its application, and it was only used to facilitate the link between the modal mass m_i and the physical masses of the configuration. According to equation (3.12) the motion of the i -th modal co-ordinate q_i is governed by :

$$m_i \ddot{q}_i(t) + k_i q_i(t) = \phi_i^T f(t)$$

Here, the “modal lever” concept will be extended by considering also the effect of excitation forces that act on the physical DOF. Interpreting q_i and $\ddot{q}_i(t)$ as the displacement and acceleration at a distance “1” from the rotation point of the lever, and locating the modal mass m_i and modal stiffness k_i at that position, one sees that the equation of motion is represented, provided that each component f_k of the force vector \mathbf{f} is placed at the corresponding DOF on the lever (Fig.3.7). The scalar product $\phi_{ik} f_k$ of each force component with the corresponding element of the mode-shape vector can thus be seen as a moment (force*distance) that acts on the modal lever, or as an equivalent force that acts at the location of q_i on the lever.

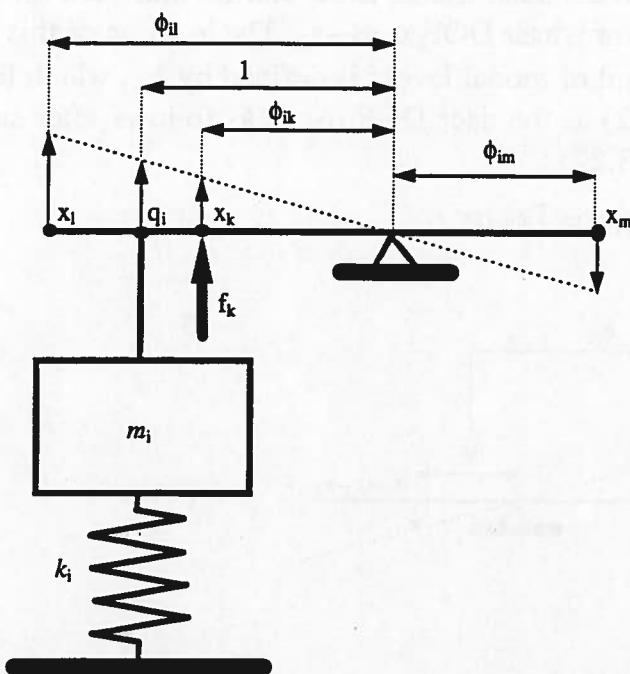


Figure 3.7 Graphical representation of mode i , including the proper location of a force component f_k , that acts on physical DOF x_k

On the basis of this graphical representation (Fig.3.7) it is not so difficult to understand the contribution of mode i to the transfer function (x_i/f_k). According to equation (3.15) this contribution equals :

$$\left(\frac{x_i}{f_k} \right)_i = \frac{\phi_{ik} \phi_{il}}{m_i s^2 + k_i}$$

Hence, the force f_k must be multiplied by the distance ϕ_{ik} in order to find the equivalent excitation force at the location of q_i on the lever, whereas the resulting modal displacement q_i must be multiplied by the distance ϕ_{il} in order to obtain the displacement of the physical DOF x_i .

3.2.3. User-Defined DOF

Often, one is not directly interested in the response of one single physical DOF, but rather in some linear combination of DOF (for example the relative displacement between a tool and a work-piece, or the relative displacement between the focus spot of a compact disc module and the track on the disk, which involves all components of the optical path of the laser light). This linear combination of physical DOF, which will be called “User-Defined DOF” or “User DOF” x_u , can be written as³ :

$$x_u = b_1 x_1 + b_2 x_2 + \dots + b_n x_n = b^T x \quad (3.22)$$

Similar to a normal physical DOF such a user DOF can be indicated on the modal lever, as illustrated in Fig.3.8 for a user DOF $x_u=x_3-x_2$. The location of this user DOF x_u with respect to the pivot point of modal level i is defined by ϕ_{iu} , which follows the same transformation rule (3.22) as the user DOF itself, as follows after substitution of equation (3.7) in equation (3.22) :

$$x_u = b^T x = b^T \sum \phi_i q_i = \sum b^T \phi_i q_i = \sum \phi_{iu} q_i \quad (3.23)$$

$$\phi_{iu} = b^T \phi_i \quad (3.24)$$

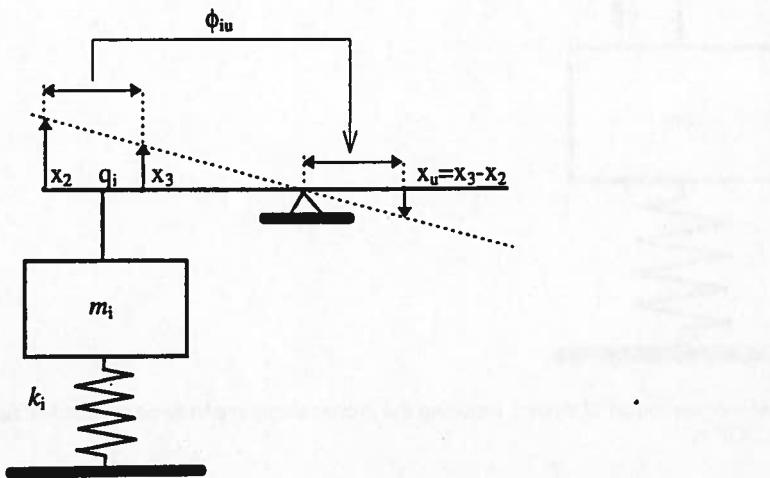


Figure 3.8 Graphical representation of mode including user DOF $x_u=x_3-x_2$

The dimension of a mode vector found in a numerical analysis is equal to the number of DOF of the simulation model, which can contain several hundreds or thousands of DOF. In an experimental modal analysis, the dimension can theoretically even be infinite, for each point on a structure displays a different behaviour. Still, from a servo-application point of view, only three user DOF are really important (all other DOF are only important for a better understanding of the nature of a mode).

³ In the case of a one-dimensional model, the definition of user DOF is straightforward, and normally does not cause any trouble. However, if applied to two- or three-dimensional structures, special attention must be paid to ensure integrity of the user DOF formula (see Appendix F).

These three user DOF define :

- input (how much a mode is excited by the servo force)
- measured output (displacement that is measured by the position sensor)
- actual output (displacement that determines the accuracy of the machine)

To illustrate this idea, the schematic representation of a servo-controlled positioning-device according to Fig.3.9 will be used. The actual task of this device is to move a payload with respect to a tool that is mounted to the machine frame. Consequently, the actual accuracy of the machine (indicated by $x_{accuracy}$) is determined by the relative motion of these two components. However, direct measurement of the distance between the tool and the payload is not possible and therefore the control action is based on the measured distance (indicated by x_{servo}) between a sensor and the slide on which the payload is mounted. The slide is driven by a linear motor which transforms the output of the controller into a force on the slide and a reaction force on the stator. It is evident that for the excitation of a mode only the result of these two forces is relevant. On the basis of either the graphical representation of Fig.3.7 or the mathematical formulation (3.13) it is evident that the resulting excitation level of a mode is determined by the relative motion x_{force} of the slide with respect to the stator.

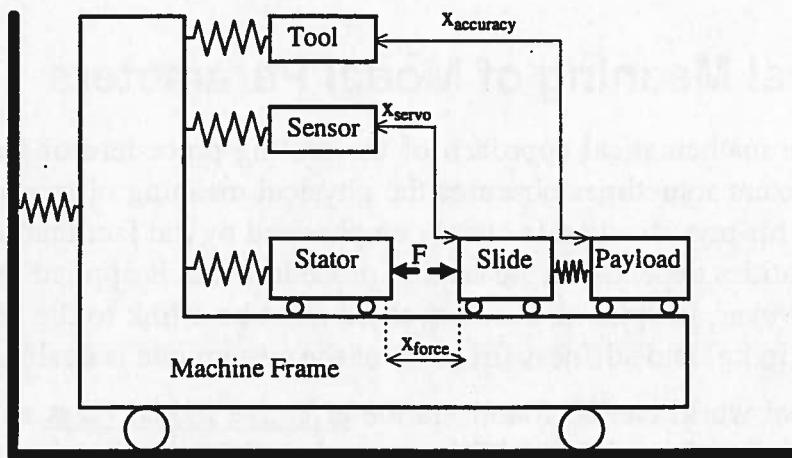


Figure 3.9 Schematic representation of servo system

Therefore, as long as one is not interested in the details of a mode the dynamics of a mode can completely be described by only three user DOF (Fig.3.10) :

- x_{servo} (example : $x_{slide} - x_{sensor}$)
- $x_{accuracy}$ (example : $x_{payload} - x_{tool}$)
- x_{force} (example : $x_{slide} - x_{stator}$)

which can easily be constructed from the contributions of the physical DOF as already illustrated in Fig.3.8.

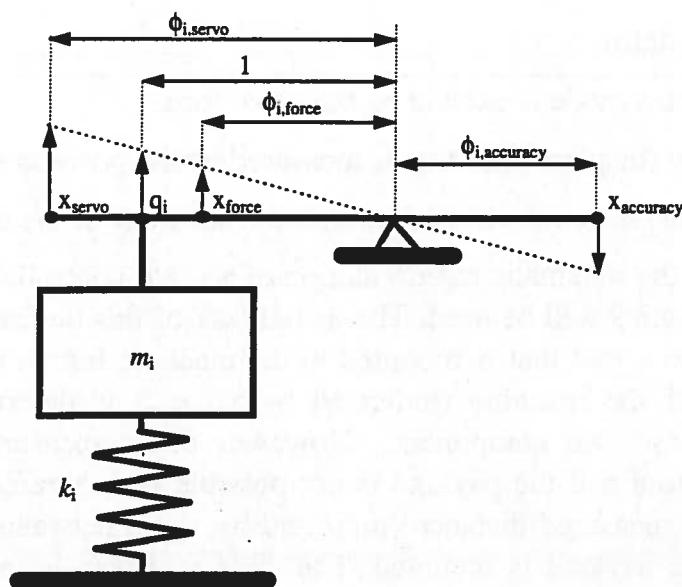


Figure 3.10 Graphical representation of the three most important user DOF of a servo system

Basically, the three values ($\phi_{i,\text{servo}}$, $\phi_{i,\text{accuracy}}$, $\phi_{i,\text{force}}$) plus the modal mass, stiffness, and damping, are sufficient to describe the contribution of mode i to the dynamics of the servo system, both in terms of frequency response and time response.

3.3. Physical Meaning of Modal Parameters

Unfortunately, the mathematical approach of the scaling procedure of mode-shapes and modal parameters sometimes obscures the physical meaning of modal mass and modal stiffness. This mystification is clearly emphasised by the fact that the value of each of these quantities depends on the scaling procedure that is applied to the mode-shape vector. However, in spite of this fact, there must be a link to the “real world” i.e. the real mass (in kg) and stiffness (in N/m) of the system one is dealing with.

The link to the real world can be found via the effective modal mass and effective modal stiffness of a mode as it is “felt” in a certain point and direction (=DOF) of the structure. These quantities are unique, do not depend on the scaling procedure, and have physical meaning and physical units.

Independent of the way in which the mode-shape vectors are normalised, one can always find an equivalent representation based on the effective modal mass and stiffness as “felt” in a certain DOF (for example x_k), as shown in Fig.3.11.

The effective modal parameters of mode i in physical DOF k can be derived from the modal parameters via the following equation, which is based on energy equivalence :

$$\begin{aligned} m_{\text{eff},ik} &= m_i / \phi_{ik}^2 \\ k_{\text{eff},ik} &= k_i / \phi_{ik}^2 \end{aligned} \quad (3.25)$$

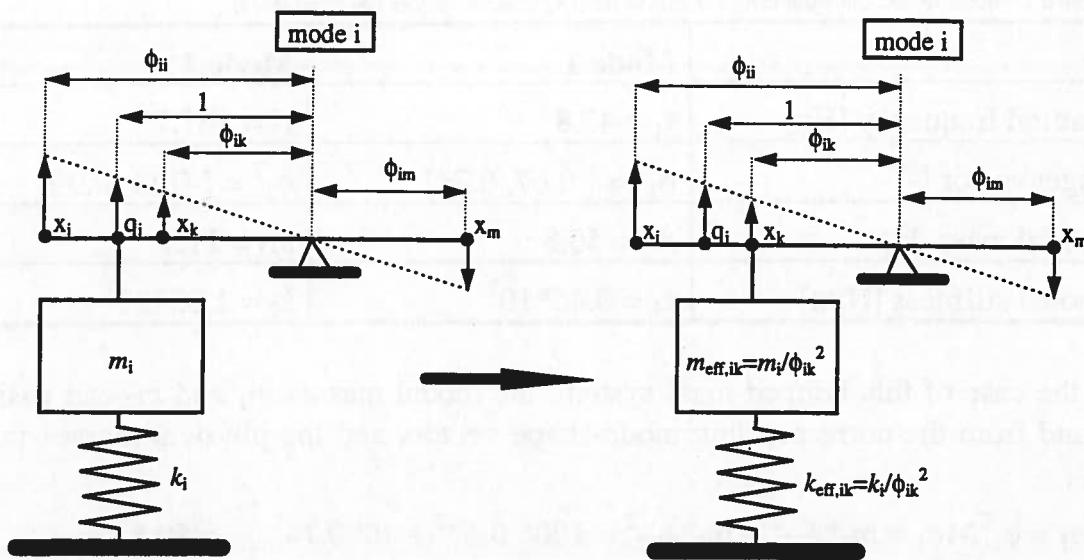


Figure 3.11 Transformation of modal parameters into effective modal values in a physical DOF (x_k), which have physical meaning and physical units

These effective modal parameters can be used very effectively in understanding topics such as sensitivity analysis (Section 3.4), or constructing the frequency response (x_k/f_k) of a complex system from the knowledge of modal contributions.

Example 1

A simple example of a two-mass spring system (Fig.3.12) will be used to illustrate the approach. The eigenvalue analysis of this system leads to the modal results given in Table 3.1, and which are graphically represented in Fig.3.13.

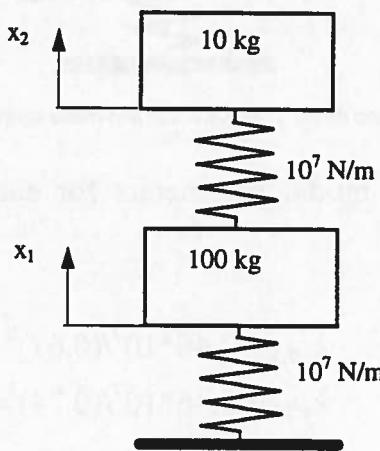


Figure 3.12 Two-mass spring system

Table 3.1 Modal results corresponding to Fig.3.12 (scaling of eigenvectors based on $\|\phi\|=1$)

	Mode 1	Mode 2
natural frequency [Hz]	$f_1 = 47.8$	$f_2 = 167.7$
eigenvector [-]	$\phi_1^T = [0.67, 0.74]$	$\phi_2^T = [-0.11, 0.99]$
modal mass [kg]	$m_1 = 50.8$	$m_2 = 11.1$
modal stiffness [N/m]	$k_1 = 0.46 \cdot 10^7$	$k_2 = 1.23 \cdot 10^7$

In the case of this lumped-mass system, the modal masses m_1 and m_2 can easily be found from the corresponding mode-shape vectors and the physical masses m_1 and m_2 :

$$m_1 = \phi_1^T M \phi_1 = m_1 * \phi_{11}^2 + m_2 * \phi_{12}^2 = 100 * 0.67^2 + 10 * 0.74^2 = 50.8 \text{ kg}$$

$$m_2 = \phi_2^T M \phi_2 = m_1 * \phi_{21}^2 + m_2 * \phi_{22}^2 = 100 * (-0.11)^2 + 10 * 0.99^2 = 11.1 \text{ kg}$$

whereas the modal stiffnesses follow from $k_i = \omega_i^2 m_i$.

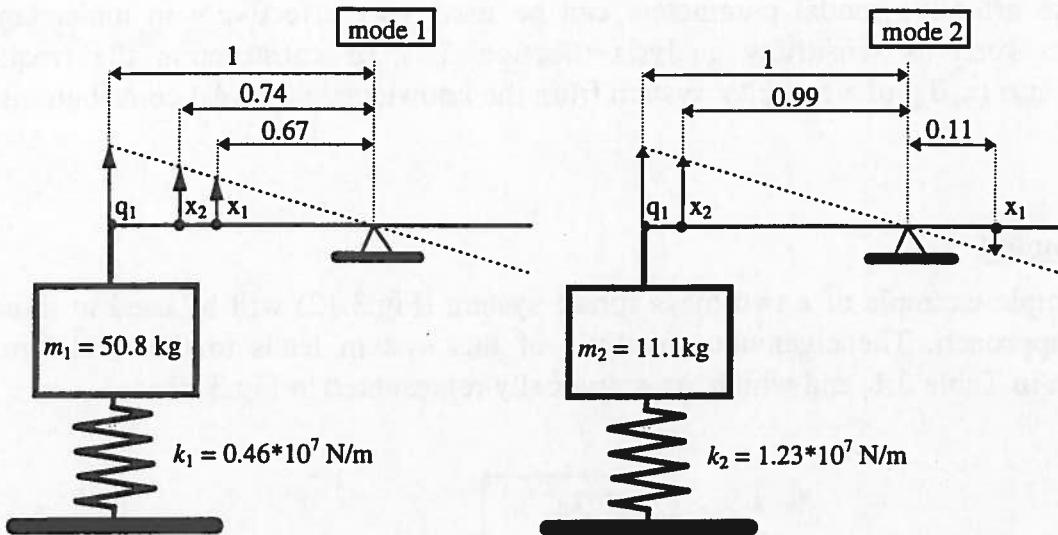


Figure 3.13 Graphical representation of modes and modal parameters of two-mass spring system

From these results, the effective modal parameters for each mode, and for each individual DOF, can be derived :

mode 1

$$m_{\text{eff},11} = 50.8 / (0.67)^2 = 112.1 \text{ kg} \quad k_{\text{eff},11} = 0.46 \cdot 10^7 / (0.67)^2 = 1.02 \cdot 10^7 \text{ N/m}$$

$$m_{\text{eff},12} = 50.8 / (0.74)^2 = 92.8 \text{ kg} \quad k_{\text{eff},12} = 0.46 \cdot 10^7 / (0.74)^2 = 0.84 \cdot 10^7 \text{ N/m}$$

mode 2

$$m_{\text{eff},21} = 11.1 / (-0.11)^2 = 927.9 \text{ kg} \quad k_{\text{eff},21} = 1.23 \cdot 10^7 / (-0.11)^2 = 101.65 \cdot 10^7 \text{ N/m}$$

$$m_{\text{eff},22} = 11.1 / (0.99)^2 = 11.2 \text{ kg} \quad k_{\text{eff},22} = 1.23 \cdot 10^7 / (0.99)^2 = 1.25 \cdot 10^7 \text{ N/m}$$

As an alternative to Fig.3.13, which is based on the modal mass and stiffness, one can for example use the effective modal parameters in DOF x_1 in the graphical representation instead (Fig.3.14). Based on this representation it is now very easy to construct the individual modal contributions to the frequency response function (x_1/F_1) of the example system, and from that to find the overall response (Fig.3.15).

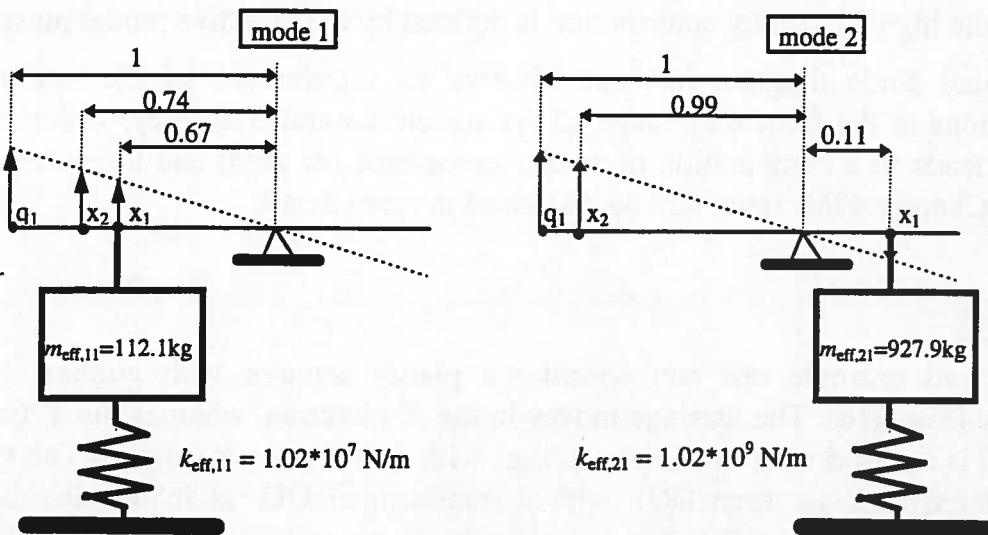


Figure 3.14 Alternative graphical representation of modes of two-mass spring system based on the effective modal mass and stiffness in DOF x_1 ,

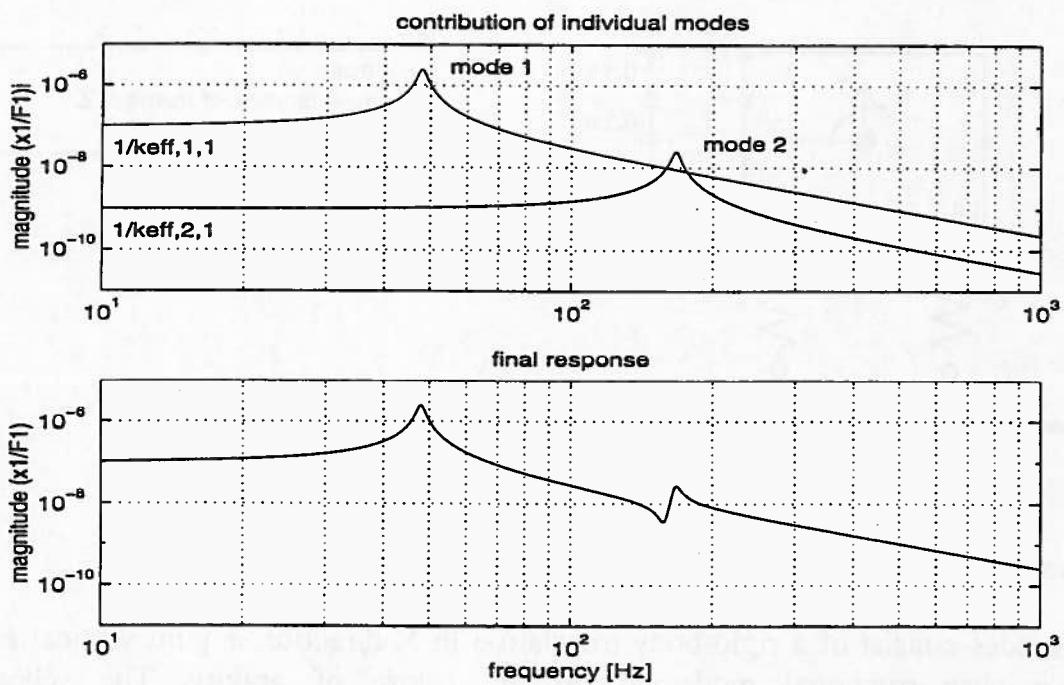


Figure 3.15 Frequency response function x_1/F_1 ,

Concentrating on the modal contributions first, one can observe that indeed the low-frequency part of each modal contribution (x_1/F_1)_i corresponds to the inverse of the calculated effective modal stiffness at DOF x_1 :

$$(x_1/F_1)_1 = 1/k_{\text{eff},11} \approx 1e^{-7} [\text{m/N}]$$

$$(x_1/F_1)_2 = 1/k_{\text{eff},21} \approx 1e^{-9} [\text{m/N}]$$

whereas the high-frequency contribution is defined by the effective modal mass.

In the final Bode diagram one can observe an interference of the two modal contributions in the frequency range of the second natural frequency, which in this example leads to a combination of an anti-resonance (or zero) and a resonance (or pole). In Chapter 4 this issue will be discussed in more detail.

Example 2

As a second example one can consider a planar actuator with guiding system flexibility (Fig.3.16). The carriage moves in the X direction, whereas the Y fixation to the rail is realised with two linear springs with stiffness c per support. The system is considered to have three DOF with a translational DOF x in the direction of movement, a translational DOF y perpendicular to the rail, and a rotational DOF ϕ . The centre of mass Z is located symmetrically between the two supports. Therefore the DOF of this system are completely decoupled if only small rotations are considered.

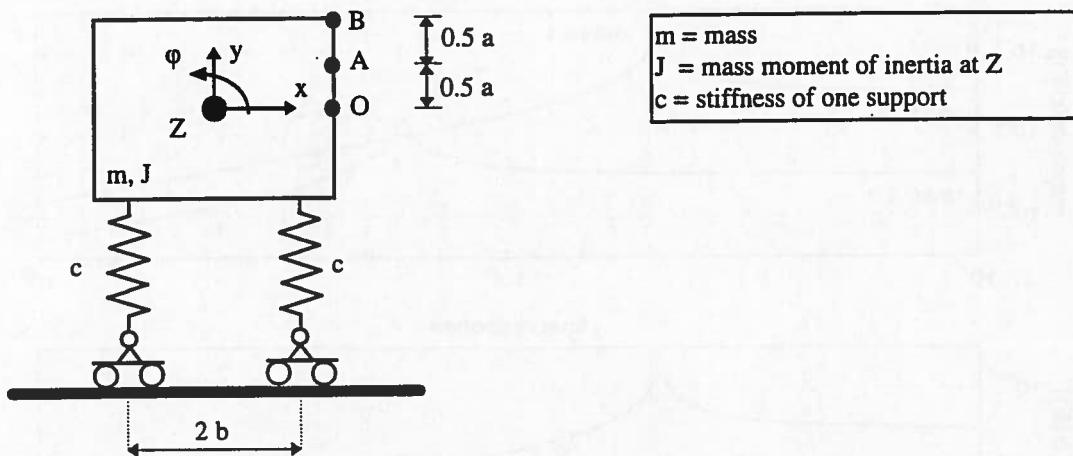


Figure 3.16 2D model of wafer chuck

The modes consist of a rigid-body translation in X direction, a pure vertical mode, and a pure rotational mode around the centre of gravity. The schematic representation of these three modes and the corresponding X motion of the points Z, O, A, and B, plus the Y motion of point Z, are indicated in Fig.3.17.

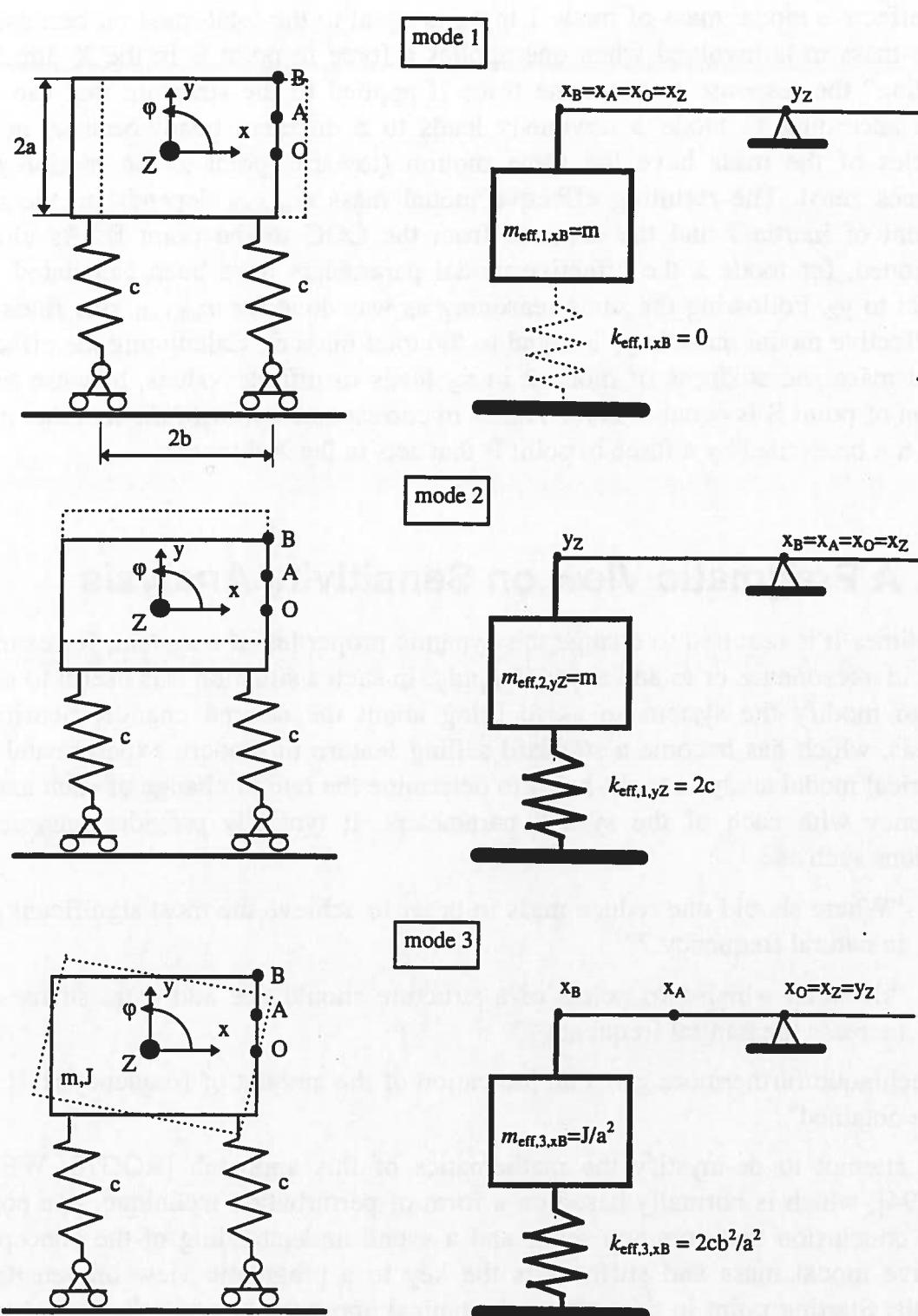


Figure 3.17 Modes of wafer chuck model

The effective modal masses and stiffnesses of modes 1 and 3 are referenced to the X motion of point B. For mode 2 the effective modal mass and stiffness is related to the Y motion of point Z, because at this mode there is no X displacement.

The effective modal mass of mode 1 in x_B is equal to the total mass m , because the entire mass m is involved when one applies a force in point B in the X direction. “Feeling” the response to the same force if applied to the structure that can only move according to mode 3 obviously leads to a different result because not all particles of the mass have the same motion (towards point Z the motion even becomes zero). The resulting effective modal mass $m_{\text{eff},3,x_B}$ depends on the mass moment of inertia J and the distance from the COG to the point B. As already mentioned, for mode 2 the effective modal parameters have been calculated with respect to y_Z . Following the same reasoning as was done for $m_{\text{eff},1,x_B}$, one finds that the effective modal mass in y_Z is equal to the total mass m . Calculating the effective modal mass and stiffness of mode 2 in x_B leads to infinite values, because the X motion of point B is equal to zero. This is in correspondence with the fact that mode 2 can not be excited by a force in point B that acts in the X direction.

3.4. A Pragmatic View on Sensitivity Analysis

Sometimes it is required to change the dynamic properties of a system, for example to avoid a resonance or to add more damping. In such a situation it is useful to know how to modify the system so as to bring about the desired change. Sensitivity analysis, which has become a standard selling feature in modern experimental and numerical modal analysis tools, helps to determine the rate of change of each natural frequency with each of the system parameters. It typically provides answers to questions such as :

- “Where should one reduce mass in order to achieve the most significant gain in natural frequency ?”
- “Between which two points of a structure should one add extra stiffness to increase the natural frequency?”

The technique furthermore gives an indication of the amount of frequency shift that can be obtained⁴.

In an attempt to de-mystify the mathematics of this approach [ROG70, WEI88, WAN94], which is normally based on a form of perturbation technique, one comes to the conclusion that common sense and a sound understanding of the concept of effective modal mass and stiffness is the key to a pragmatic view on sensitivity analysis. Starting point in this non-mathematical approach is again the modal lever representation of a mode (Fig.3.18).

⁴ It should be noted that the predicted changes of the natural frequencies, which are derived on the basis of a sensitivity analysis, are only accurate for small parameter modifications.

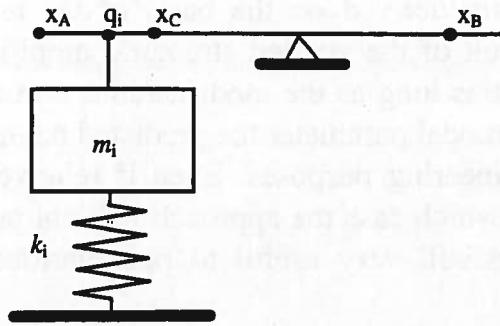


Figure 3.18 Graphical Representation of a mode with 3 DOF (x_A , x_B , x_C)

Assuming that one is asked to increase the natural frequency of this mode by attaching a linear spring k between any two DOF (for simplicity only three physical DOF are indicated on the lever) it does not take an expensive tool or a genius mind to come up with the notion that the relative motion between x_A and x_B is the largest of all possible combinations, and thus the effectiveness of an additional spring k between these two DOF is maximal (maximum contribution to potential energy).

Very much in line with this concept is the idea of considering the kinetic and potential energy distribution of a system over its inertias and connecting flexibilities. For every modal deformation pattern, the relative contribution of the various components and connectors to the modal kinetic and potential energy can be calculated. If one has to increase the frequency of a mode, one should focus on stiffening those components or connectors with the highest contribution to the modal potential energy. On the other hand, components that contribute significantly to the modal kinetic energy are serious candidates for mass reduction.

A first-order approximation of the new natural frequency of mode i can easily be derived by considering the effective modal mass and stiffness of that mode in the relevant DOF. In the case of an extra mass Δm in DOF x_k , the effective modal mass $m_{\text{eff},i}$ in that DOF is required, whereas in the case of an additional spring Δk between two DOF x_k and x_l one has to compare the contribution of Δk to the effective modal stiffness $k_{\text{eff},i}$ in the user DOF (x_k-x_l). The new natural frequency of mode i will be approximately :

extra mass Δm ⁵

$$f_{\text{new},i} = \frac{1}{2\pi} \sqrt{\frac{k_{\text{eff},i}}{m_{\text{eff},i} + \Delta m}} = \frac{1}{2\pi} \sqrt{\frac{k_{\text{eff},i}}{m_{\text{eff},i}}} \sqrt{\frac{m_{\text{eff},i}}{m_{\text{eff},i} + \Delta m}} = f_{\text{old},i} \sqrt{\frac{m_{\text{eff},i}}{m_{\text{eff},i} + \Delta m}} \quad (3.26)$$

extra stiffness Δk

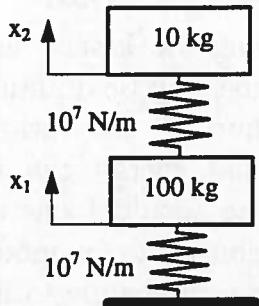
$$f_{\text{new},i} = \frac{1}{2\pi} \sqrt{\frac{k_{\text{eff},i} + \Delta k}{m_{\text{eff},i}}} = \frac{1}{2\pi} \sqrt{\frac{k_{\text{eff},i}}{m_{\text{eff},i}}} \sqrt{\frac{k_{\text{eff},i} + \Delta k}{k_{\text{eff},i}}} = f_{\text{old},i} \sqrt{\frac{k_{\text{eff},i} + \Delta k}{k_{\text{eff},i}}} \quad (3.27)$$

⁵ in case of a mass reduction, Δm will be a negative number

Equations (3.26) and (3.27) are derived on the basis of the assumption that the modes do not change as a result of the applied structural modification. In general, this assumption is not true, but as long as the modifications remain small compared to the corresponding effective modal parameter the predicted natural frequencies will be accurately enough for engineering purposes. Even if relatively large structural modifications are involved, in which case the approach will not predict the modified eigenfrequencies reliably, it is still very useful to rank various proposed design modifications.

Example 1

The simple example of a two-mass spring system (Fig.3.19), which was already discussed in Section 3.3, will be used to illustrate the approach. First, the frequency shift due to an extra mass at DOF x_2 will be investigated, then the influence of an additional spring between x_2 and x_1 will be shown.



	mode 1	mode 2
eigenfrequency	$f_1 = 47.8 \text{ Hz}$	$f_2 = 167.7 \text{ Hz}$
modal mass	$m_1 = 50.8 \text{ kg}$	$m_2 = 11.1 \text{ kg}$
modal stiffness	$k_1 = 0.46 \cdot 10^7 \text{ N/m}$	$k_2 = 1.23 \cdot 10^7 \text{ N/m}$
eigenvector	$\phi_1^T = [0.67, 0.74]$	$\phi_2^T = [-0.11, 0.99]$

Figure 3.19 Two-mass spring system and the result of a modal analysis (scaling of eigenvectors based on $|\phi_i|=1$)

Extra mass Δm at x_2

In order to analyse the effect of an extra mass at x_2 , the effective modal mass at that DOF needs to be known for both modes. Recalling the results from Section 3.3 :

$$\text{mode 1} : m_{\text{eff},12} = 92.8 \text{ kg}$$

$$\text{mode 2} : m_{\text{eff},22} = 11.2 \text{ kg}$$

one can estimate the new natural frequencies on the basis of equation (3.26) and compare the results to the exact results of an eigenvalue analysis of the modified system :

extra mass $\Delta m=1 \text{ kg}$ added to m_2

$$f_1 = 47.8 * \sqrt{92.8 / 93.8} = 47.5 \text{ Hz} \quad (\text{exact result: } 47.5 \text{ Hz})$$

$$f_2 = 167.7 * \sqrt{11.2 / 12.2} = 160.7 \text{ Hz} \quad (\text{exact result: } 160.7 \text{ Hz})$$

extra mass $\Delta m=10 \text{ kg}$ added to m_2

$$f_1 = 47.8 * \sqrt{92.8/102.8} = 45.4 \text{ Hz} \quad (\text{exact result: } 45.2 \text{ Hz})$$

$$f_2 = 167.7 * \sqrt{11.2/21.2} = 121.9 \text{ Hz} \quad (\text{exact result: } 125.2 \text{ Hz})$$

Extra stiffness between x_2 and x_1

To estimate the influence of extra stiffness between the two DOF, one needs to calculate the effective modal stiffness that corresponds to the relative motion between x_2 and x_1 . As already explained (Section 3.2.3) the concept of modal levers can be extended to any linear combination of physical DOF. Therefore, one can extend the graphical representation of the two modes with the relative motion (x_2-x_1), as indicated in Fig.3.20.

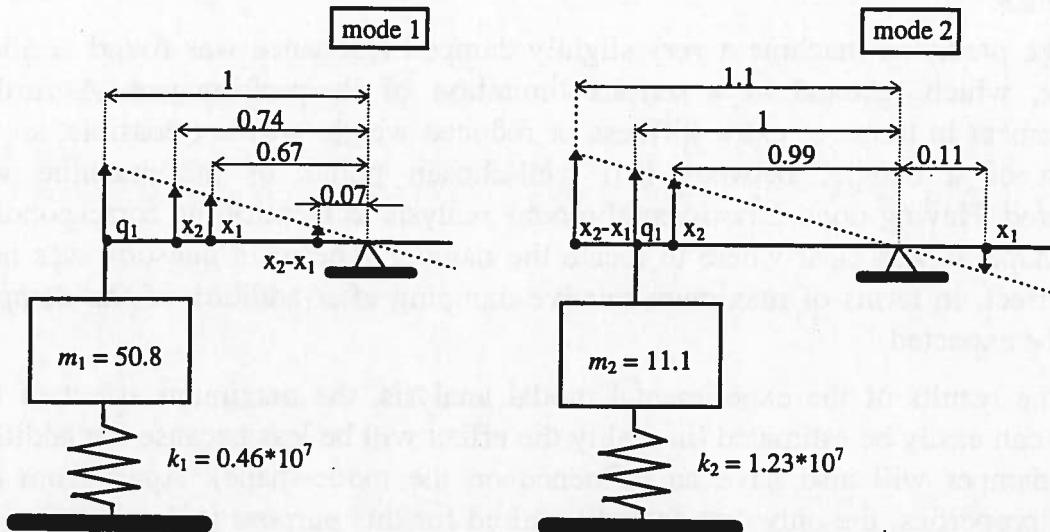


Figure 3.20 Graphical representation of modes and modal parameters of two-mass spring system

From this graphical representation it is now very easy to derive the effective modal stiffness at the newly defined user DOF. This is done in precisely the same way as for any ordinary physical DOF :

$$\text{mode 1 : } k_{\text{eff},1(2-1)} = 0.46 * 10^7 / (0.07)^2 = 93.9 * 10^7 \text{ N/m}$$

$$\text{mode 2 : } k_{\text{eff},2(2-1)} = 1.23 * 10^7 / (1.1)^2 = 1.0 * 10^7 \text{ N/m}$$

On the basis of these values and equation (3.27) the effect of additional stiffness between x_2 and x_1 will be estimated and compared to the exact results :

extra stiffness $\Delta k=10^7 \text{ N/m}$ between x_2 and x_1

$$f_1 = 47.8 * \sqrt{94.9/93.9} = 48.1 \text{ Hz} \quad (\text{exact result: } 47.9 \text{ Hz})$$

$$f_2 = 167.7 * \sqrt{2/1} = 237.2 \text{ Hz} \quad (\text{exact result: } 236.6 \text{ Hz})$$

extra stiffness $\Delta k = 10^8$ N/m between x_2 and x_1

$$f_1 = 47.8 * \sqrt{103.9 / 93.9} = 50.3 \text{ Hz} \quad (\text{exact result: } 48.0 \text{ Hz})$$

$$f_2 = 167.7 * \sqrt{11 / 1} = 556.2 \text{ Hz} \quad (\text{exact result: } 553.9 \text{ Hz})$$

As can be seen from this example, the very pragmatic approach to sensitivity analysis without too much emphasis on the mathematical treatment yields correct first order approximations of the new natural frequencies after a modification. As it is only an approximation and not a complete new eigenvalue analysis, the results are better for small changes as compared to the effective modal parameters. Yet, even if the modifications are significant, the approach provides a quick scan of the effect in terms of frequency shift that can be expected from a certain design modification.

Example 2

In a high precision machine a very slightly damped resonance was found at about 150 Hz, which resulted in a serious limitation of its performance. A further improvement in terms of extra stiffness or reduced weight was not possible, so the addition of a damper between two well-chosen points of the machine was considered. Having done experimental modal analysis to identify the corresponding mode shape, it was clear where to locate the damper. The main question was how much effect, in terms of maximum relative damping after addition of the damper, was to be expected.

Using the results of the experimental modal analysis, the maximum effect of the damper can easily be estimated (in reality the effect will be less because the addition of the damper will also have an influence on the mode-shape). Apart from the damper properties, the only data that is required for this purpose is the modal mass and stiffness, plus the modal displacements of the two points (x_2 and x_1) between which the damper will be added (Fig.3.21).

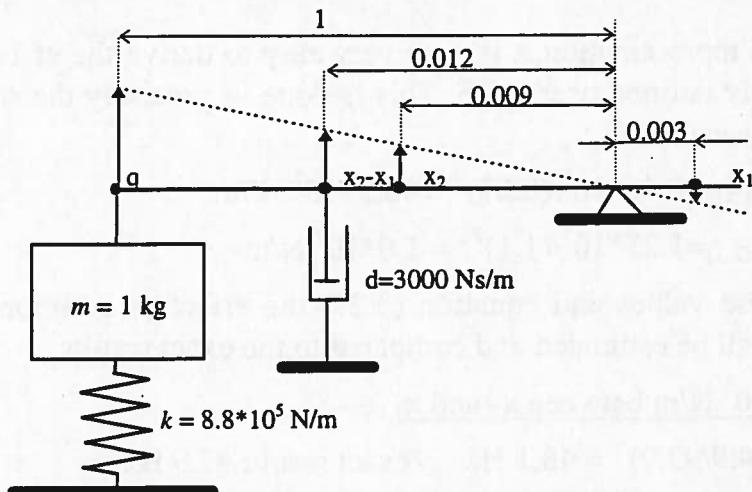


Figure 3.21 Modal data of 150 Hz mode for the estimation of the effect of a damper between DOF x_2 and x_1

The relative damping of the mode as a result of the additional damper d can be approximated on the basis of the effective modal properties in the user-defined DOF ($x_2 - x_1$), which corresponds to the relative displacement between the damper connection points. The effective modal mass and stiffness in the user DOF ($x_2 - x_1$) equal :

$$m_{\text{eff}} = 1/(0.012)^2 = 6944 \text{ kg}$$

$$k_{\text{eff}} = 8.8 * 10^5 / (0.012)^2 = 6.1 * 10^9 \text{ N/m}$$

Using these values, the resulting relative modal damping β can be found :

$$\beta = \frac{d}{2\sqrt{m_{\text{eff}} k_{\text{eff}}}} = \frac{3000}{2\sqrt{6944 * 6.1 * 10^9}} = 0.15\%$$

This value is disappointingly low, because typical relative damping values for a machine in which no special damping measures have been taken are about 0.1-2%. The main reason for this disappointing low value is the huge effective mass and stiffness that the damper "feels" in the user DOF ($x_2 - x_1$).

3.5. Modal Superposition

Previously, the lever representation of individual mode-shapes was introduced in order to describe the basic idea in a physically oriented way. Similarly, the idea of modal superposition can be visualised by utilising an approach for adding up motion, which was already used by Koster in the dynamic analysis of CAM mechanisms [KOS73]. In the mechanism shown in Fig.3.22, the motion of the output y equals the sum of the motions of the two inputs x_1 and x_2 . Setting x_1 equal to 0 one finds $y=x_2$, equally $x_2=0$ renders $y=x_1$ and if both contribution are non-zero, one finds $y=x_1+x_2$. Addition of more than two contributions can be achieved by placing several of these basic mechanisms in series.

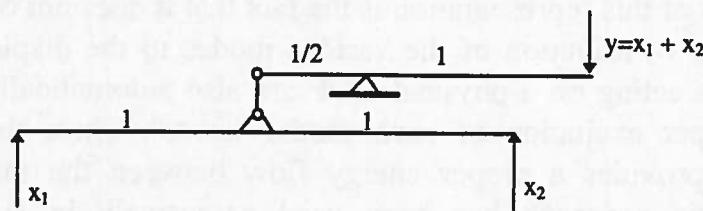


Figure 3.22 Addition of motion

This approach can also be applied to the concept of modal superposition, which expresses the motion of any physical DOF $x_k(t)$ as a summation of modal contributions according to equation (3.17) :

$$x_k(t) = \sum_{i=1}^n \phi_{ik} q_i(t) = \sum_{i=1}^n x_{ki}(t)$$

Where the motion of a physical DOF x_k is derived from only two modal contributions x_{ki} and x_{kj} , (this is not a fundamental choice), one can basically copy Fig.3.22 and replace the names of the variables, as shown in Fig.3.23.

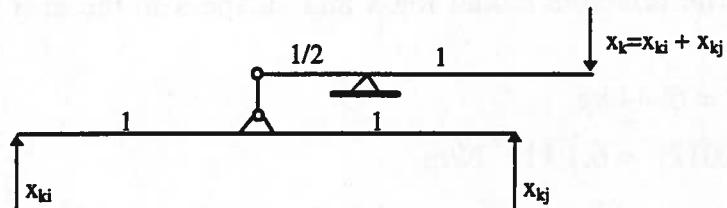


Figure 3.23 Addition of modal contributions

Combining this concept of summation of modal contributions with the lever representation of mode-shapes finally leads to Fig.3.24, which is believed to be a very powerful visualisation for the transformation between the modal and the physical domains (again for simplicity only two modes are taken into account).

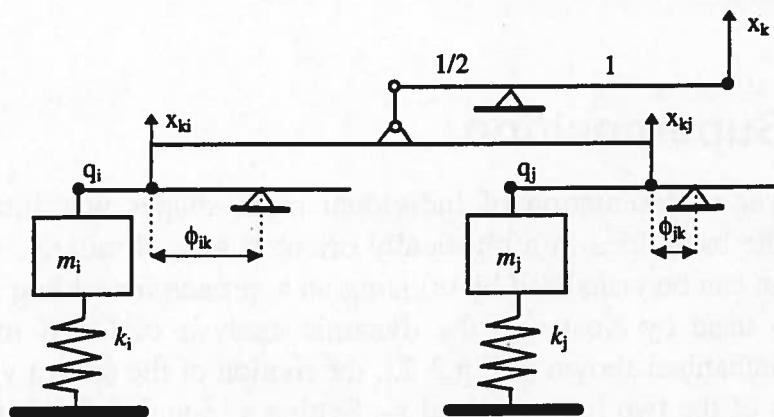


Figure 3.24 Conversion between Modal DOF and Physical DOF

The interesting aspect of this representation is the fact that it does not only provide a proper addition of the contribution of the various modes to the displacement of a physical DOF. Forces acting on a physical DOF are also automatically transmitted and provide the proper excitation of each modal "lever". Thus, this physically oriented conversion provides a proper energy flow between the modal and the physical domain. This approach has been used extensively in simulations of mechatronic devices which have been carried out in various electronic network analysis programs [RAN88/4].

3.6. Suspension Modes

The fact that the automotive and aeronautic industries have played a dominant role in the development of the concept of modal analysis is also reflected in some of the terminology that is being used in modal studies.

A rocket or an aeroplane in flight will possess rigid body modes which are entirely determined by the mass properties, and in which there is absolutely no deformation of the structure. The natural frequency of such a mode is equal to 0 Hz.

Unfortunately, this terminology is also used when considering the lowest natural frequencies of a machine or product that are caused by the flexibility of the suspension system. In such a situation the term "rigid-body mode" misleadingly suggests that the structure exhibits no internal deformation. Even renowned textbooks on modal testing [EWI84] contribute to this misunderstanding by maintaining the terminology even when considering softly supported test structures, "so that the rigid-body modes, while no longer having zero natural frequencies, have values which are very low in relation to those of the bending modes", without pointing out the implications. A better term for such a mode would be "suspension mode", because it expresses the main characteristics of the mode without the misleading implication that the machine behaves as a rigid body.

To illustrate the importance of the internal deformation a very simplified physical model of a precision machine, such as an electron microscope, will be considered (see Fig.3.25).

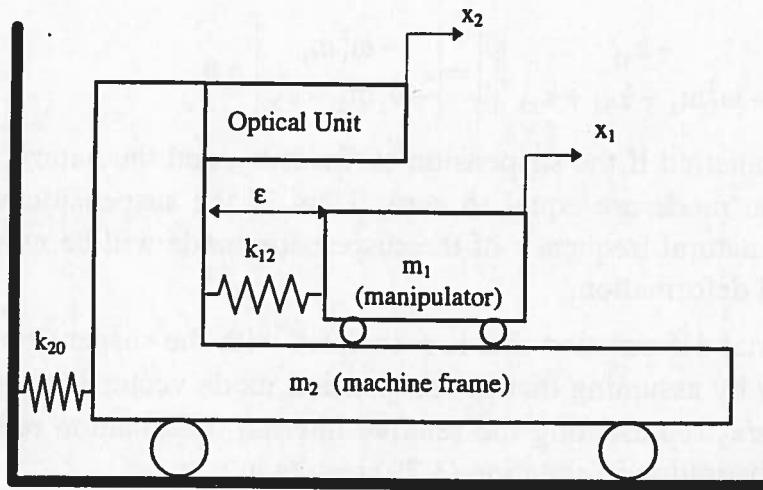


Figure 3.25 Simplified physical model of a precision machine

The machine basically consists of a very heavy granite machine frame to which an optical unit is rigidly connected (combined mass equals m_2). The optical unit takes images of a specimen that is mounted on a manipulator (mass m_1) that has a certain internal flexibility with respect to the granite machine frame. For a proper operation

of the machine, the internal deformation $\varepsilon = x_2 - x_1$ needs to be minimal. Typically, such a machine is designed for a high internal stiffness and it is furthermore very softly supported in order to prevent external (floor) vibrations from entering the machine and endangering its performance.

Assuming that the natural frequency ω_1 of the suspension mode ϕ_1 is significantly lower than that of the internal mode, one can approximate the frequency of the suspension mode by considering the motion of the entire machine as one rigid body on the stiffness of the suspension system. This approach evidently leads to a correct estimate of the suspension frequency, but one should keep in mind that there is always a small amount of internal deformation in case of a non-zero suspension stiffness k_{20} . This can be demonstrated quite straightforwardly by considering the unforced equations of motion of the system :

$$\begin{bmatrix} m_1 & 0 \\ 0 & m_2 \end{bmatrix} \begin{bmatrix} \ddot{x}_1 \\ \ddot{x}_2 \end{bmatrix} + \begin{bmatrix} k_{12} & -k_{12} \\ -k_{12} & k_{20} + k_{12} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0 \quad (3.28)$$

or, written as an eigenvalue problem,

$$\begin{bmatrix} -\omega^2 m_1 + k_{12} & -k_{12} \\ -k_{12} & -\omega^2 m_2 + k_{20} + k_{12} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0 \quad (3.29)$$

Both modes must satisfy equation (3.29). Assuming that the suspension mode is free of internal deformation, the modal displacement of x_1 and x_2 would be identical (and consequently $\phi_1^T = [1, 1]^T$ ⁶). Substituting this vector in the left-hand side of equation (3.29) leads to :

$$\begin{bmatrix} -\omega_1^2 m_1 + k_{12} & -k_{12} \\ -k_{12} & -\omega_1^2 m_2 + k_{20} + k_{12} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -\omega_1^2 m_1 \\ -\omega_1^2 m_2 + k_{20} \end{bmatrix} = 0 \quad (3.30)$$

which can only be satisfied if the suspension stiffness k_{20} and the natural frequency ω_1 of the suspension mode are equal to zero. Thus, if the suspension stiffness is unequal to zero, the natural frequency of the suspension mode will be non-zero, and there will be internal deformation.

The amount of internal deformation that is associated with the suspension mode can be found very easily by assuming that the suspension mode vector is equal to $\phi_1^T = [1 + \varepsilon_r, 1]^T$, with $\varepsilon_r = \varepsilon/x_2$ representing the relative internal deformation referenced to the motion of x_2 . Substitution in equation (3.29) results in :

$$\begin{bmatrix} -\omega_1^2 m_1 + k_{12} & -k_{12} \\ -k_{12} & -\omega_1^2 m_2 + k_{20} + k_{12} \end{bmatrix} \begin{bmatrix} 1 + \varepsilon_r \\ 1 \end{bmatrix} = \begin{bmatrix} -\omega_1^2 m_1 + k_{12} \varepsilon_r \\ -\omega_1^2 m_2 + k_{20} - k_{12} \varepsilon_r \end{bmatrix} = 0 \quad (3.31)$$

⁶ Only the direction, but not the length, of the mode-shape vector is uniquely defined, therefore a convenient choice has been made (see also Chapter 3.1)

The internal deformation associated with the suspension mode can now be found from the first part of equation (3.31) and amounts to :

$$\varepsilon = \varepsilon_r x_2 = \omega_1^2 \frac{m_1}{k_{12}} x_2 = \frac{\omega_1^2}{\omega_{int}^2} x_2; \quad \omega_{int} = \sqrt{\frac{k_{12}}{m_1}} \quad (3.32)$$

with ω_{int} representing the natural frequency of the manipulator where the baseframe is clamped or infinitely heavy.

Equation (3.32) expresses that the internal deformation associated with the suspension mode depends on the ratio of the natural frequencies of suspension mode and internal mode. The larger the natural frequency of the internal mode compared to the suspension mode, the smaller the internal deformation. The equation also confirms the previous statement that the internal deformation associated with the suspension mode can only be zero when the natural frequency of that mode is zero.

It is worthwhile mentioning that equation (3.32) can also be derived in a more intuitive way by considering the internal deformation as a quasi-static deformation due to the acceleration forces imposed on the manipulator during a vibration of the entire machine at the suspension frequency ω_1 . During such a vibration an acceleration force needs to be transferred from the base frame to the manipulator in order to make the manipulator follow the motion x_2 of the base frame. The resulting internal deformation is found from this force and the internal stiffness k_{12} :

$$F = m_1 \ddot{x}_2 = m_1 \omega_1^2 x_2 \quad (3.33)$$

$$\varepsilon = \frac{F}{k_{12}} = \frac{m_1 \omega_1^2}{k_{12}} x_2 = \frac{\omega_1^2}{\omega_{int}^2} x_2 \quad (3.34)$$

As one can see, this engineering approach yields the same results as found previously (equation (3.32)).

As an example of a situation in which the internal deformation associated with the suspension mode is of significant importance, one could consider a high-precision machine, such as an electron microscope that is excited, for example due to floor vibrations, such that it vibrates on its suspension with an amplitude of $100\mu\text{m}$ and a frequency of 3Hz . Assuming that the internal frequency of the manipulator is equal to 150Hz , the internal deformation of the machine can be calculated from equation (3.32) :

$$\varepsilon = \frac{3^2}{150^2} 100\mu\text{m} = 40\text{nm}$$

which seems to be negligibly small, but not in high-precision applications, such as wafer-steppers or electron microscopes, in which one is aiming at resolutions in the nanometer range.

I would say that every education has three important dimensions: the first is the knowledge of the world around us; the second is the development of our own personal qualities; and the third is the development of our ability to contribute to the world.

Education is not just about learning facts and figures, it's also about developing skills and abilities, and learning how to think critically.

Education is not just about learning facts and figures, it's also about developing skills and abilities, and learning how to think critically.

Education is not just about learning facts and figures, it's also about developing skills and abilities, and learning how to think critically.

Education is not just about learning facts and figures, it's also about developing skills and abilities, and learning how to think critically.

Education is not just about learning facts and figures, it's also about developing skills and abilities, and learning how to think critically.

Education is not just about learning facts and figures, it's also about developing skills and abilities, and learning how to think critically.

Education is not just about learning facts and figures, it's also about developing skills and abilities, and learning how to think critically.

Education is not just about learning facts and figures, it's also about developing skills and abilities, and learning how to think critically.

Education is not just about learning facts and figures, it's also about developing skills and abilities, and learning how to think critically.

Education is not just about learning facts and figures, it's also about developing skills and abilities, and learning how to think critically.

Education is not just about learning facts and figures, it's also about developing skills and abilities, and learning how to think critically.

Education is not just about learning facts and figures, it's also about developing skills and abilities, and learning how to think critically.

4. Modes & Servo Stability

In Section 2.3 the effect of machine dynamics on the performance of positioning devices was introduced. One of the two limiting effects mentioned, namely the destabilisation of the servo control loop and, as a consequence of this, the limitation of bandwidth and disturbance rejection, will be discussed in this chapter.

The interaction between the desired (rigid body) motion and the dynamics of one additional mode and its effect on the frequency response function ($x_{\text{servo}}/F_{\text{servo}}$) and the stability of the control loop is the basis of this chapter. It will be linked to the three major dynamic phenomena, that were already introduced in Section 2.4, and design guidelines will be presented. Finally, a general approach that can be used to limit the destabilising effect of modes is discussed and illustrated by an example.

This chapter focuses completely on industrial applications, which are generally based on standard, fixed-order, motion controllers. Commonly, a PID feedback section is implemented in these controllers. The more elaborated versions sometimes have extra low-pass and notch filtering facilities, and sometimes a feedforward section, as described in Chapter 2, is available. The limitations of a fixed order controller structure imply that control solutions based on inverse plant models are not considered here. Rather, it is the mechanical system that has to be designed in such a way that the required bandwidth and disturbance rejection can be achieved without running the risk of servo stability problems.

4.1. Basic Characteristics of Mechanical FRF

As a starting point for the discussion of the destabilising effect of dynamics on a position control loop, such as shown in Fig.4.1, the mechanical FRF ($x_{\text{servo}}/F_{\text{servo}}$) will be investigated.

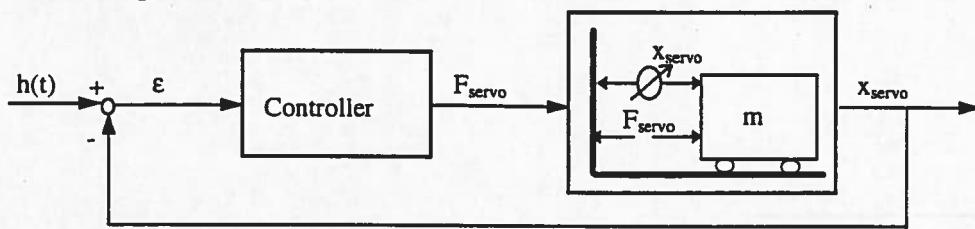


Figure 4.1 Mechanical position servo-system

In the ideal situation the mechanical system behaves as one rigid body with mass m , so the mechanical transfer function ($x_{\text{servo}}/F_{\text{servo}}$) can be written as:

$$\left(\frac{x_{\text{servo}}}{F_{\text{servo}}} \right) = \frac{1}{ms^2} \quad (4.1)$$

The corresponding Bode and Nyquist representations are shown in Fig.4.2. In the entire frequency range, the amplitude diagram shows a -2 slope, and the phase equals -180° . In the Nyquist diagram the curve travels the negative real axis, starting at $-\infty$ and approaching 0 for increasing frequencies.

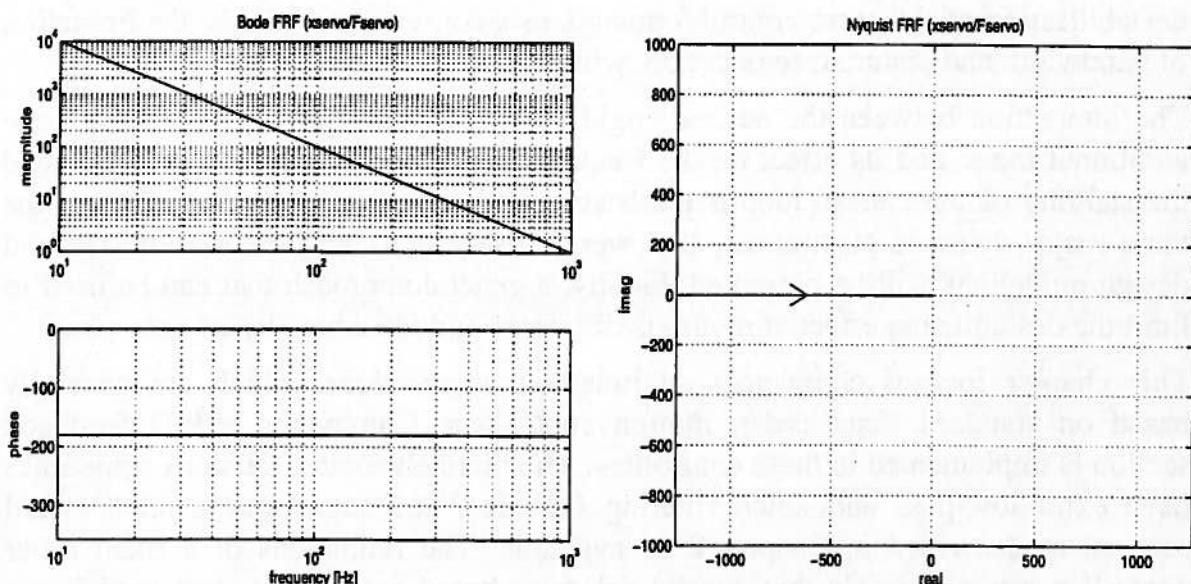


Figure 4.2 FRF ($x_{\text{servo}}/F_{\text{servo}}$) of an ideal system with no resonances (a. Bode, b. Nyquist representation)

In the case of one extra modal contribution¹, the equation for the mechanical transfer function needs to be extended with one extra term, which can be derived from the graphical representation of a mode (Fig.4.3) or from equation (3.15).

$$\left(\frac{x_{\text{servo}}}{F_{\text{servo}}} \right) = \frac{1}{ms^2} + \frac{\phi_{i,\text{servo}} \phi_{i,\text{force}}}{m_i s^2 + k_i} = \frac{1}{ms^2} + \frac{\phi_{i,\text{servo}} \phi_{i,\text{force}}}{m_i s^2 + \omega_i^2 m_i} \quad (4.2)$$

¹ The additional mode is assumed to have a small amount of modal damping (typically 1-2%). However, for reasons of clarity, this damping term is not included in the equations.

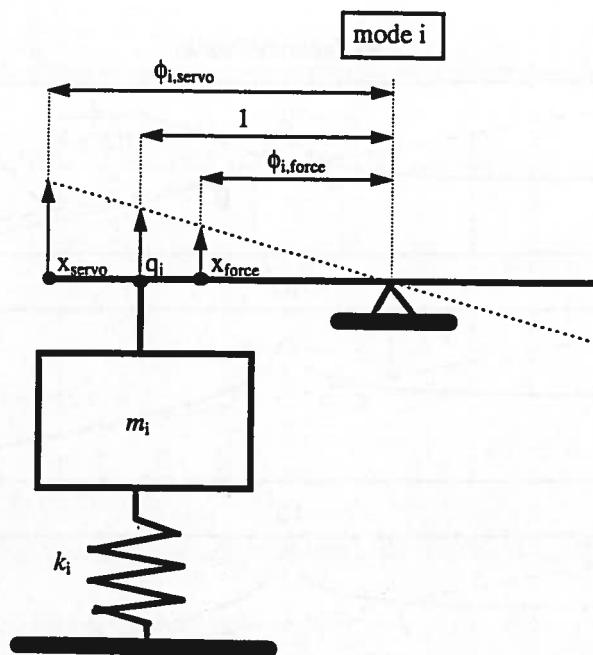


Figure 4.3 Graphical representation of mode *i*

The final transfer function and the exact interaction between the two parts in equation (4.2) depends on the values of the various parameters. However, one can only distinguish six different combinations which lead to four different interaction patterns in the final FRF. To derive these six combinations, equation (4.2) is reduced to its essentials by introducing the variable α , which relates the high-frequency contribution of the mode to that of the rigid-body motion :

$$\alpha = \frac{\left(\frac{\phi_{i,servo} \phi_{i,force}}{m_i} \right)}{\left(\frac{1}{m} \right)} = \frac{\phi_{i,servo} \phi_{i,force}}{m_i} m \quad (4.3)$$

which simplifies equation (4.2) to :

$$\left(\frac{x_{servo}}{F_{servo}} \right) = \frac{1}{ms^2} + \frac{\alpha}{ms^2 + m\omega_i^2} \quad (4.4)$$

Equation (4.4) will be the basis for the discussion of the various patterns that can be observed in frequency response functions and of the effect of resonances on servo stability. It is important to note that equation (4.4) is not a special case but a general description of the behaviour of a mechanical servo-system that contains a resonance.

Investigating equation (4.4), and considering the general shape of the two contributions in a Bode diagram (Fig.4.4), reveals that apart from the precise locations of the natural frequency ω_i , three different types of intersection pattern can be found in the amplitude plot.

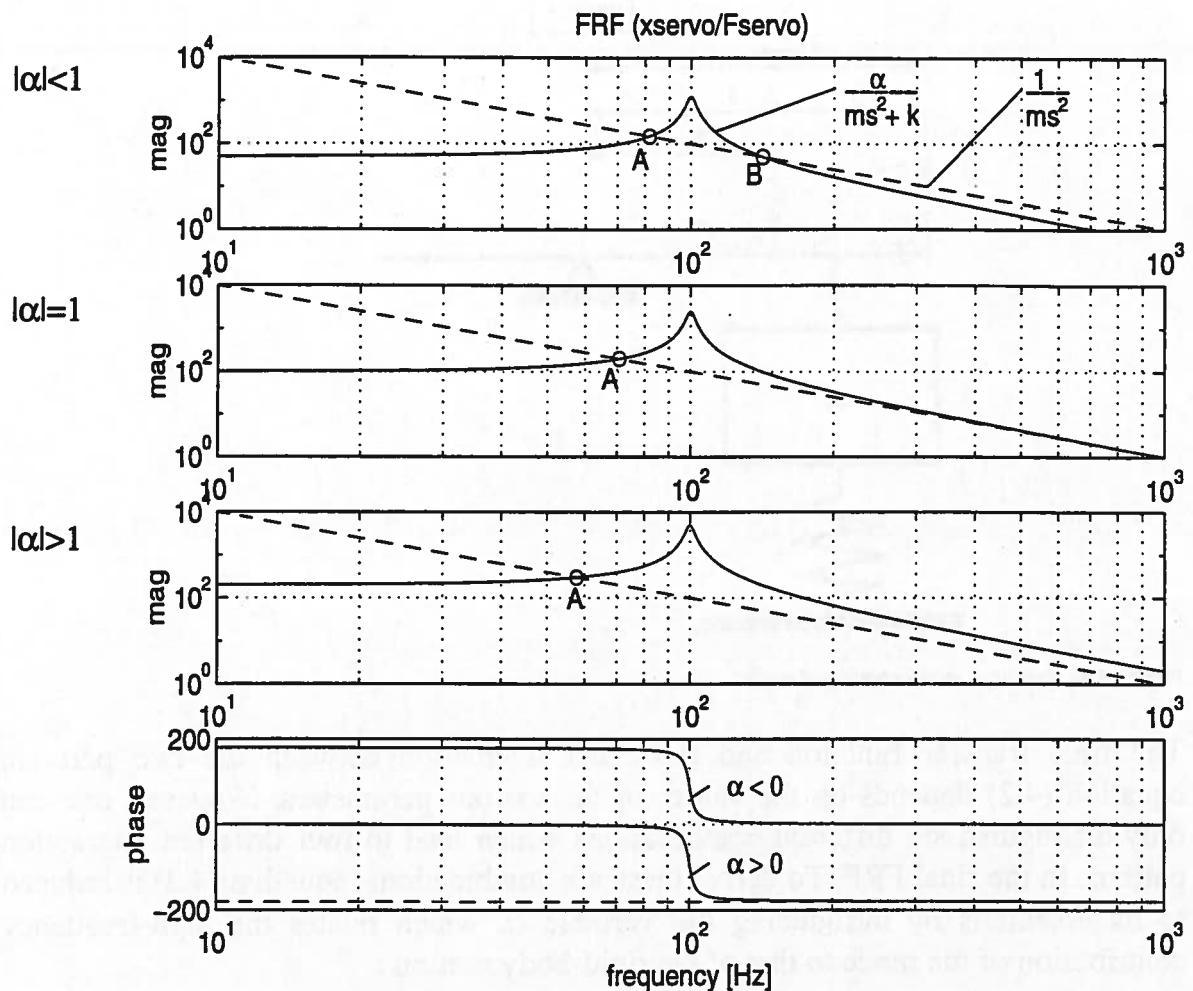


Figure 4.4 Contribution of rigid-body motion and modal dynamics to the amplitude and phase of FRF ($x_{\text{servo}}/F_{\text{servo}}$) for various values of α

Depending on the absolute values of α , one can observe :

- $|\alpha| < 1$: two intersections
- $|\alpha| = 1$: one intersection and asymptotic approach at high frequencies
- $|\alpha| > 1$: one intersection

The final interaction between the rigid-body motion and the additional mode will not only depend on $|\alpha|$ but also on the sign of α^* , which determines the phase relation between the two contributions. This doubles the number of possible combinations to six.

* A positive value of α implies that for the mode under consideration the motion of the point at which the force is applied (more precisely : the relative motion of the points between which the servo force is applied) has the same sign as the motion of the measuring point (again, more precisely : the relative motion of the points between which the position is measured).

Basically, it is fairly easy to construct the general shape of the overall FRF from the magnitude and phase of the two parts. Interesting points are the interaction of the two parts at the frequency that corresponds to an intersection in the amplitude plot. At this frequency the magnitudes are equal, so it depends on the phase of the two contributions whether they cancel each other, thus leading to a zero², or just add up.

This approach can be used to interpret and understand the Bode diagrams of the resulting FRF of each of the six distinct combinations, shown in Fig.4.5.

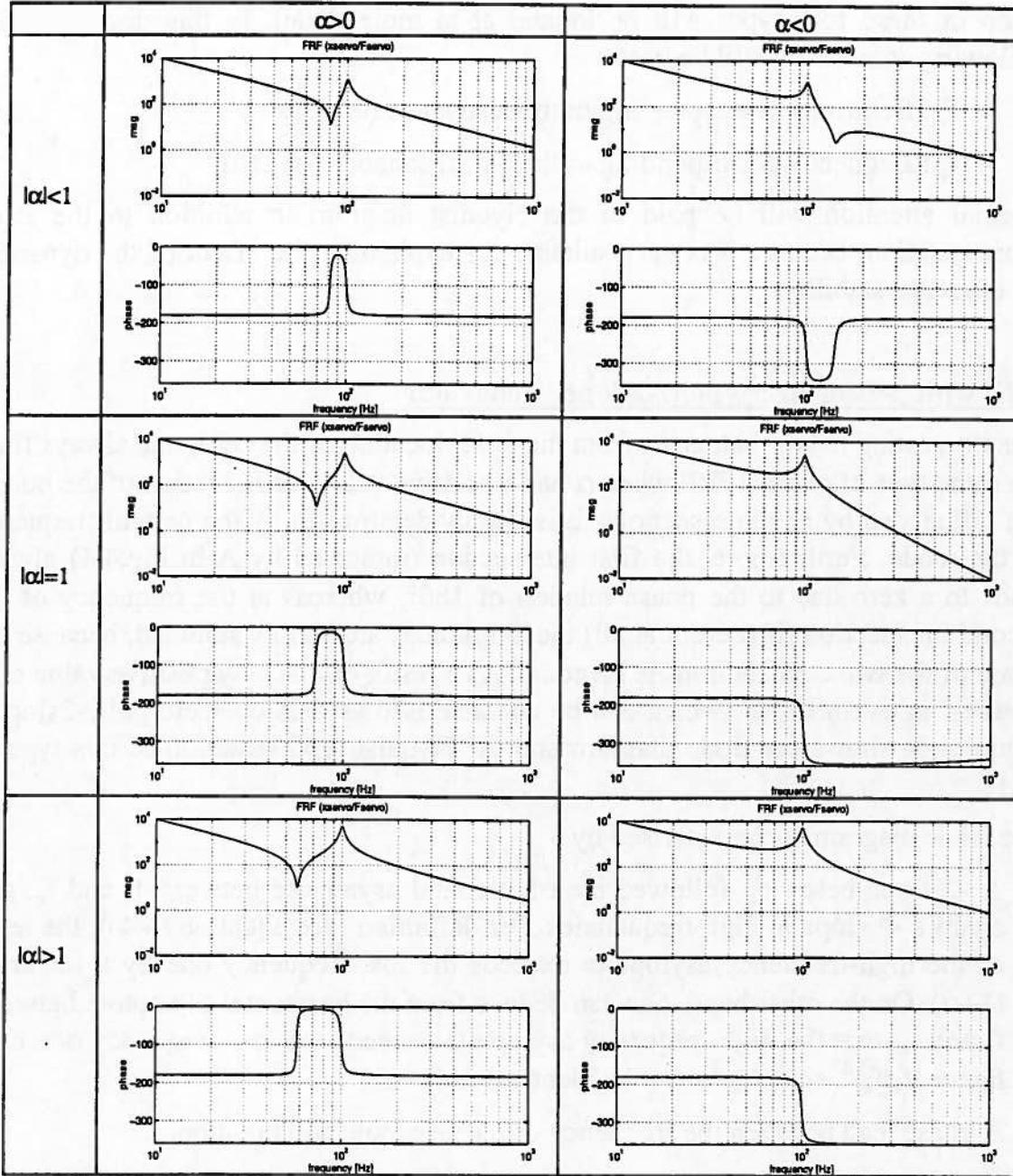


Figure 4.5 Bode diagram of final FRF ($x_{\text{servo}}/F_{\text{servo}}$) for six values of α

² A zero is also very often referred to as "anti-resonance".

When analysing the plots of Fig.4.5, four different types of FRF can be found, which can be characterised as :

- -2slope/zero/pole/-2slope ($\alpha > 0$)
- -2slope/pole/zero/-2slope ($-1 < \alpha < 0$)
- -2slope/pole/-4slope³ ($\alpha = -1$)
- -2slope/pole/-2slope ($\alpha < -1$)

Each of these four types will be looked at in more detail. In this discussion the following convention will be used :

- f_e : frequency corresponding to the resonance (or pole)
- f_a : frequency corresponding to the anti-resonance (or zero)

Special attention will be paid to the Nyquist diagram in addition to the Bode representation, because it is very suitable for explaining the effect of the dynamics on the servo stability.

FRF with “-2slope/zero/pole/-2slope” behaviour

It is interesting to see that, apart from the exact location of the zero, one always finds the same type of overall FRF when α has a positive value. The location of the pole is not influenced by any intersections, it is simply determined by the natural frequency of the mode. Furthermore, the first intersection (indicated by A in Fig.4.4) always leads to a zero due to the phase relation of 180° , whereas at the frequency of the second intersection (if present at all) the amplitudes are simply summed, because the phase of the two contributions is identical. As a result of this, any positive value of α leads to an overall FRF, which can be characterised as “-2slope/zero/pole/-2slope”. Figure 4.6a shows the Bode diagram and the Nyquist representation of this type of FRF.

The Bode diagram is characterised by :

- A -2 slope below f_a , followed by a horizontal asymptote between f_a and f_e , and again a -2 slope at high frequencies. Per definition (see equation (4.4)), the level of the high-frequency asymptote exceeds the low-frequency one by a factor of $(1+\alpha)$. On the other hand, one can deduce from the horizontal asymptote between f_a and f_e , that the high-frequency asymptote exceeds the low-frequency one by a factor $(f_e/f_a)^2$, so $(f_e/f_a)^2$ must be identical to $(1+\alpha)$.
- A phase lead between the frequency of the zero/pole combination.

³ Because of the presence of relative modal damping β the high-frequency asymptote is actually characterised by a -3 slope for frequencies above $f_e/(2\beta)$; with $\beta=1\%$ this crossover frequency equals $50*f_e$. This explains the slight increase of the phase plot in the investigated frequency range.

In the Nyquist representation, one sees that the zero/pole combination leads to a huge, almost circular, excursion, which lies to the right of the original rigid-body curve when travelling from low to high frequencies.

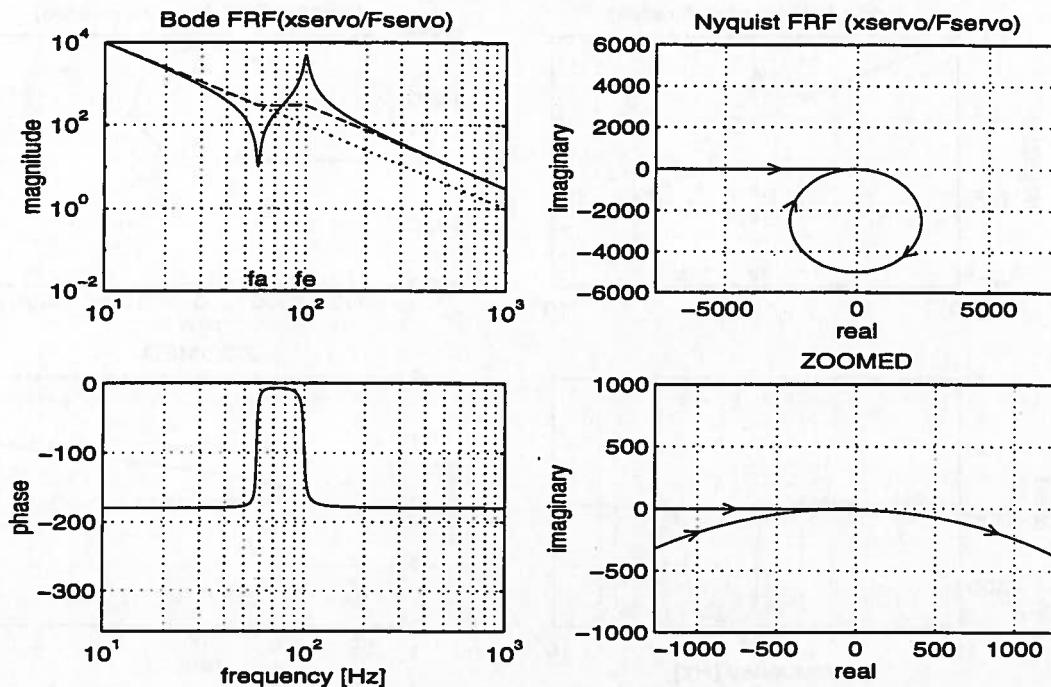


Figure 4.6a Bode and Nyquist representations of a FRF of type “-2slope/zero/pole/-2slope”

FRF with “-2slope/pole/zero/-2slope” behaviour

The combination “-2slope/pole/zero/-2slope” is only found for $-1 < \alpha < 0$. The pole is again determined by the presence of the modal resonance, whereas the zero emerges because at the second intersection (indicated by B in Fig.4.4) the amplitudes of the two parts are equal but the phases are opposite. At the first intersection (point A in Fig.4.4) no cancelling of contributions takes place because of the correspondence in phase. The Bode and Nyquist representations of this mechanical FRF are shown in Fig.4.6b.

Typical characteristics in the Bode diagram are :

- A high-frequency mass line (-2 asymptote), which is always lower than the low-frequency one, whereas between frequencies f_e and f_a the behaviour is characterised by a -4 asymptote. The level of the high-frequency asymptote differs a factor $(1+\alpha)$ from the low-level asymptote, which is identical to the ratio $(f_e/f_a)^2$.
- A phase lag between the frequency of the pole/zero combination.

In contrast to the previous FRF, the circular excursion in the Nyquist diagram now lies to the left of the original rigid-body curve when travelling from low to high frequencies, and it is not closed.

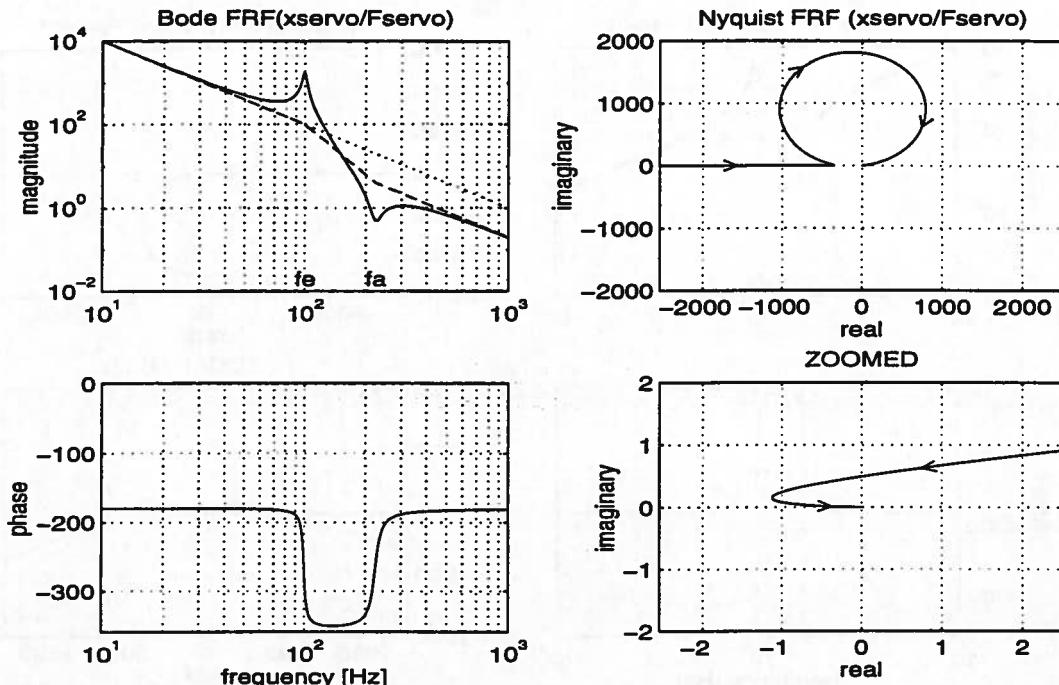


Figure 4.6b Bode and Nyquist representation of a FRF of type “-2slope/pole/zero/-2slope”

FRF with “-2slope/pole/-4slope” behaviour

The combination “-2slope/pole/-4slope” is only found for $\alpha=-1$. As it is not evident that the combination of a rigid-body motion and one extra mode with $\alpha=-1$ leads to a FRF with a -4 slope at high frequencies, one needs to go back to equation (4.4) and re-arrange it slightly. For $\alpha=-1$:

$$\left(\frac{x_{\text{servo}}}{F_{\text{servo}}} \right) = \frac{1}{ms^2} + \frac{-1}{ms^2 + m\omega_i^2} = \frac{(ms^2 + m\omega_i^2) - ms^2}{ms^2(ms^2 + m\omega_i^2)} = \frac{\omega_i^2}{ms^2(s^2 + \omega_i^2)} \quad (4.5)$$

which shows that the high-frequency asymptote is indeed a -4 slope⁴. The phase undergoes a shift from -180° to -360° at the resonance frequency.

Figure 4.6c shows the Bode diagram and the Nyquist representation of this type of FRF. Apart from the high-frequency behaviour (phase approaches -360°), the

⁴ In the presence of relative modal damping β the transferfunction is defined by :

$$\left(\frac{x_{\text{servo}}}{F_{\text{servo}}} \right) = \frac{2\beta\omega_i s + \omega_i^2}{ms^2(s^2 + \omega_i^2)}$$

which shows that the -4 slope behaviour changes into a -3 slope behaviour at $\omega=\omega_i/(2\beta)$. With relative modal damping in the range of 1% this crossover frequency lies at about $50*\omega_i$, which is far beyond the frequency range of interest and it is therefore justified to characterise the behaviour as “-2slope/pole/-4slope”.

Nyquist plot looks very similar to the previous FRF (Fig.4.6b), showing a huge excursion on the left-hand side.

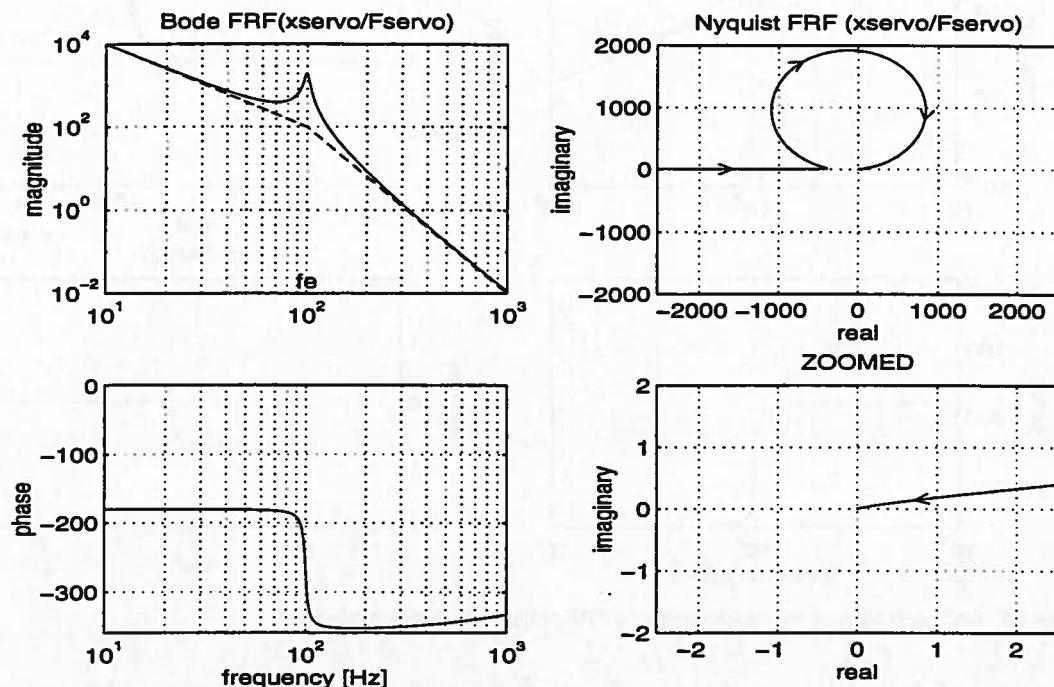


Figure 4.6c Bode and Nyquist representations of a FRF of type “-2slope/pole/-4slope”

FRF with “-2slope/pole/-2slope” behaviour

The last combination, “-2slope/pole/-2slope” is found for $\alpha < -1$. The location of the pole is again determined by the modal resonance. Only one intersection of the two curves is found in Fig.4.4 (point A) but no cancelling of contributions takes place due to the correspondence in phase. The Bode and Nyquist representations of this mechanical FRF are shown in Fig.4.6d.

Typical characteristics in the Bode diagram are :

- A -2 asymptote both in the low-frequency and the high-frequency range. The high-frequency asymptote differs by a factor $|1+\alpha|$ from the low-frequency one, which for $\alpha < -1$ leads to $0 < |1+\alpha| < \infty$. Consequently, the high-frequency asymptote can be lower or higher than the low-frequency one, depending on the value of α .
- The phase undergoes a shift from -180° to -360° at the resonance frequency.

Similar to the previous FRF, the circular excursion in the Nyquist diagram lies to the left of the original rigid-body curve when travelling from low to high frequencies, it is not closed and the point $(0,0)$ is approached with a phase of -360° .

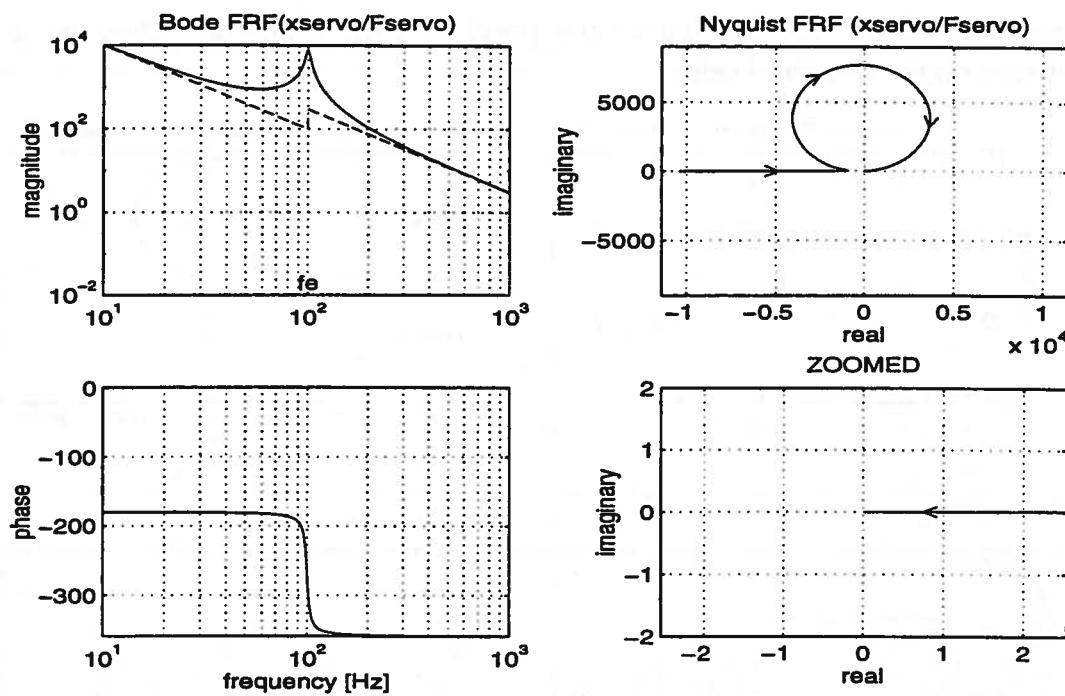


Figure 4.6d Bode and Nyquist representations of a FRF of type "2slope/pole/2slope"

This type of behaviour corresponds to a so-called "non-minimum phase system" [DOF86, FRA88]. Typical for this kind of systems is the presence of a zero in the right-hand s-plane. For the underlying system this can be demonstrated by rearranging equation (4.4) :

$$\left(\frac{x_{\text{servo}}}{F_{\text{servo}}} \right) = \frac{1}{ms^2} + \frac{\alpha}{ms^2 + m\omega_i^2} = \frac{ms^2 + m\omega_i^2 + \alpha ms^2}{ms^2(ms^2 + m\omega_i^2)} = \frac{(1 + \alpha)s^2 + \omega_i^2}{ms^2(s^2 + \omega_i^2)} \quad (4.6)$$

The zeros of transfer function (4.6) are defined by :

$$(1 + \alpha)s^2 + \omega_i^2 = 0 \quad (4.7)$$

which, for $\alpha < -1$, leads to the two real zeros $z_{1,2}(s)$ defined in expression (4.8). One of these zeros does indeed lie on the positive real axis in the s-plane, which explains the non-minimum phase behaviour.

$$z_{1,2}(s) = \pm \frac{\omega_i}{|1 + \alpha|} \quad (4.8)$$

In the time domain, a non-minimum phase system manifests itself by a step response that initially has a sign which is opposite to the sign of the input. This is not a very favourable behaviour, which should not occur in a normal mechatronic positioning-device, so this type of system will be excluded from detailed discussion in this thesis. A short reference will only be made in Section 4.3. which describes the effect of guiding-system flexibility on the mechanical FRF.

4.2. Destabilising Effect of Modes

In this section the effect of each of the three basic mechanical FRF (type “-2slope/pole/-2slope” has already been disqualified in the previous section) on the stability of the control loop will be discussed. It is impossible to give a complete coverage of each of these characteristic shapes with any type of controller. Because of its huge practical application in industrial motion-control systems, a PID controller with additional second order low-pass filter will be the basis for the discussion of stability⁵. No judgement is made regarding the quality of the control loop, because this strongly depends on the specifications (settling time, overshoot, disturbance rejection). Merely the influence of dynamics on the stability of the loop is demonstrated.

First, the open-loop response of a single mass without additional dynamics, in combination with a well-tuned controller will be looked at. Typical crossover frequencies for a PID controller with second order low-pass filtering are (relative to the bandwidth frequency f_b) :

$$f_i = f_b/10$$

$$f_d = f_b/3$$

$$f_{lp} = 4*f_b$$

which leads to the asymptotic amplitude plot shown in Fig.4.7.

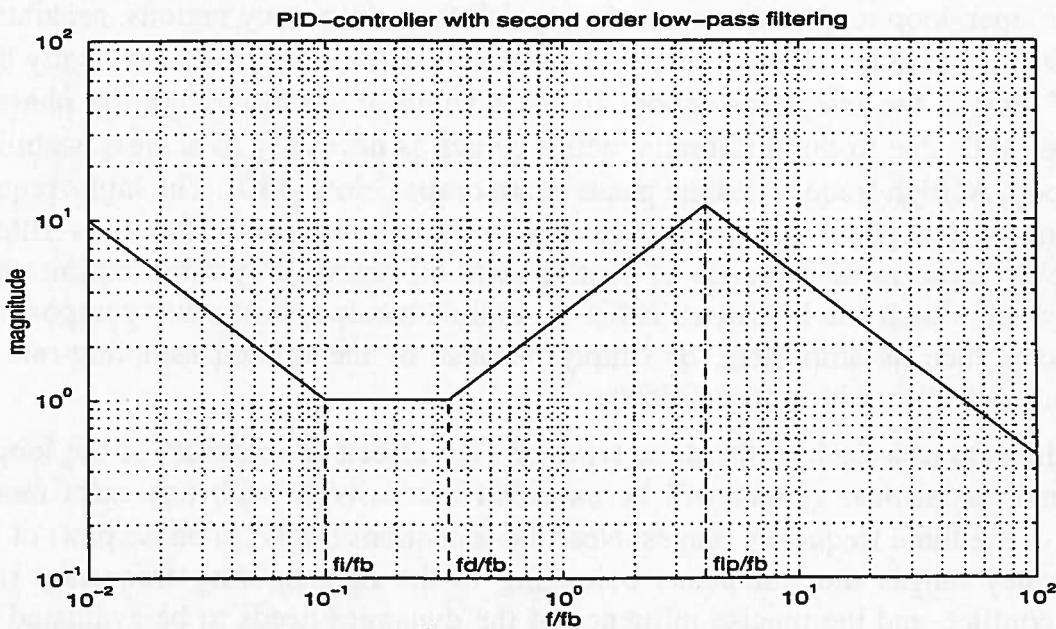


Figure 4.7 Typical crossover frequencies of a PID-controller with 2.order low-pass filtering

⁵Suppression of a resonance by means of a notch filter can sometimes be useful; however, its practical applicability depends very strongly on the reproducibility of the resonance phenomenon, both within a single system but also among several systems that have to work with identical controller settings.

With these settings, the open-loop response of the position loop (controller + mechanics) looks like Fig.4.8. As one can see, a safe amplitude margin of a factor 3 was selected, while the phase margin amounts to about 50° .

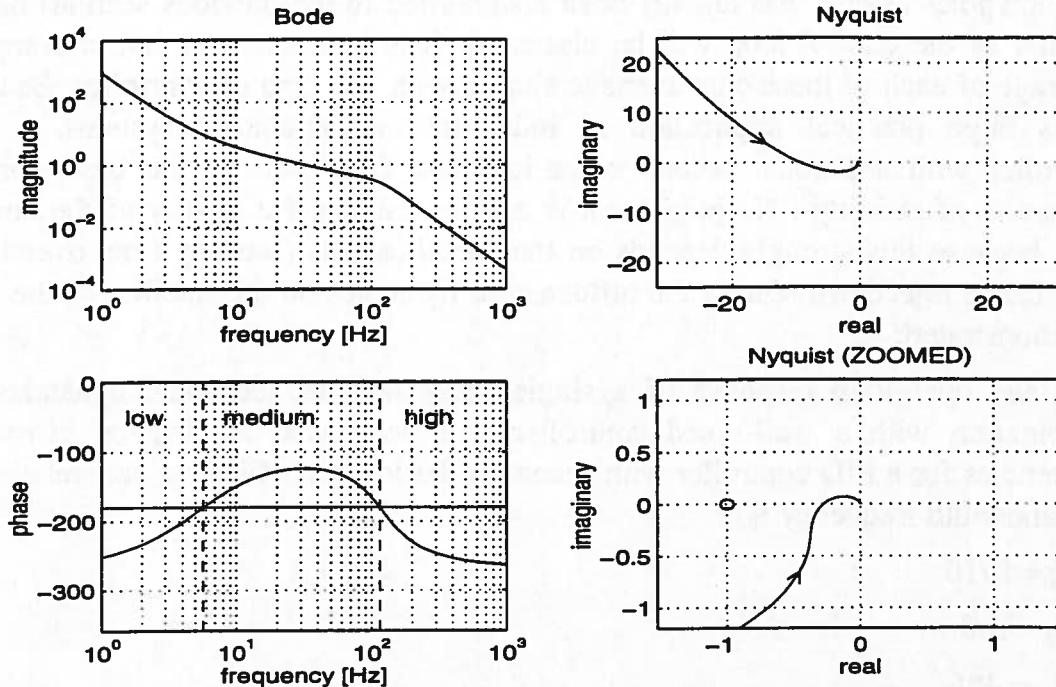


Figure 4.8 Ideal open-loop FRF of a position servo without mechanical resonances ($f_b = 30\text{Hz}$)

In the open-loop response one can distinguish three frequency regions, separated by a -180° crossing in the phase plot. At low frequencies, the phase is typically below -180° due to the integrator action. In the medium-frequency range the phase lies above -180° due to the differential action which is necessary to achieve stability of the loop. At high frequencies the phase again drops below -180° . The high-frequency decline of the phase is typically caused by the second order low-pass filtering; however, even in the absence of such a filter all practical systems exhibit similar behaviour, which can be caused either by limited bandwidth of other components in the loop such as amplifiers, or simply because of the limited sampling-rate of a digital controller, which causes delay.

To illustrate how each of the three types of FRF affects the stability of the loop, the original mechanical system will be extended successively with one extra mode in each of the three frequency ranges. Near the transitions (-180° in phase plot) of these frequency ranges the statements belonging to the neighbouring frequency ranges may conflict, and the precise influence of the dynamics needs to be evaluated from case to case.

From this discussion one can determine whether a certain resonance phenomenon has the potential to destabilise the control loop. Apart from the type of behaviour and its frequency, it also depends heavily on the amplitude of the modal contribution

whether instability really occurs. In Section 4.3 guidelines are given on how to minimise the amplitude of an undesired mode to the open-loop FRF.

4.2.1. FRF of type “-2slope/zero/pole/-2slope”

If this type of dynamics occurs in the low-frequency region⁶ (Fig.4.9a) it can have a destabilising effect (depending also on the amount of damping and the value of α), because the phase lead that is introduced results in a crossing of the -180° line, while at the same time the amplitude at the anti-resonance frequency can approach “1”.

In the medium-frequency region the zero/pole combination has no destabilising effect, as can be concluded from Fig.4.9b. No extra -180° crossing is introduced by the phase lead, and in the Nyquist plot the huge resonance excursion lies to the right of the curve, thus increasing the distance with respect to the point $(-1,0)$.

In the high-frequency range (Fig.4.9c), the phase lead results in an extra -180° crossing, which can become critical if the resonance causes the amplitude plot to approach “1”. In the Nyquist diagram the resonance bulb again lies to the right of the original curve, which in this frequency range means that it can come close to the point $(-1,0)$.

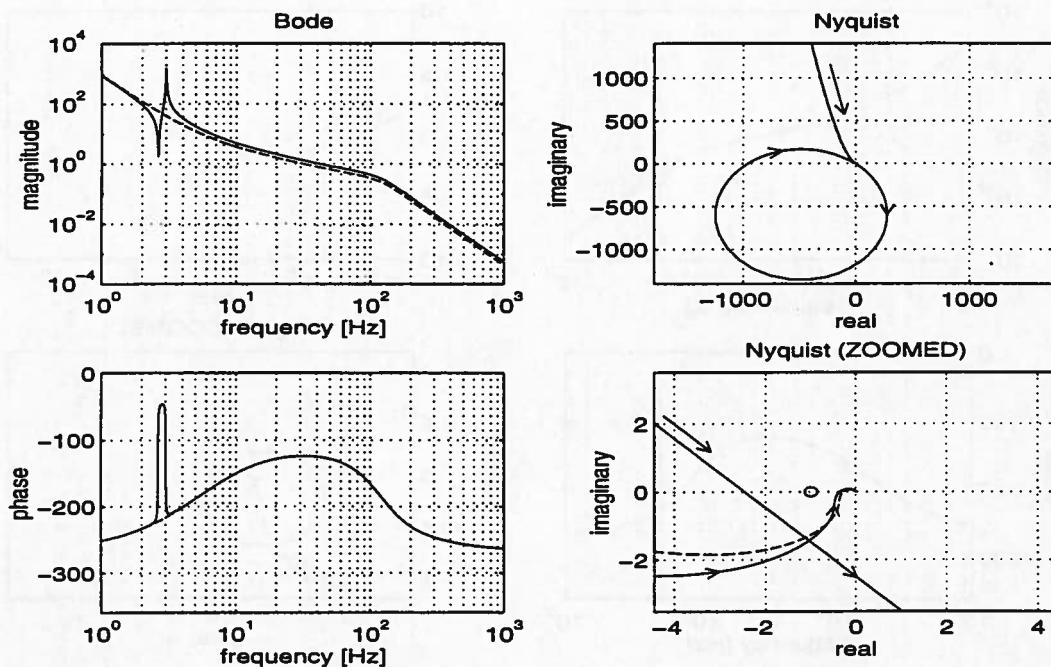


Figure 4.9a Open loop FRF of type “-2slope/zero/pole/-2slope” with low-frequency resonance

⁶ It happens only very seldom that a resonance in this frequency range has a limiting effect. In a well-designed system, the only dynamic effect in this frequency range is that of the machine frame that resonates on its fixation. This generally has an $\alpha < 0.1$ and a relative damping $\beta > 2\%$.

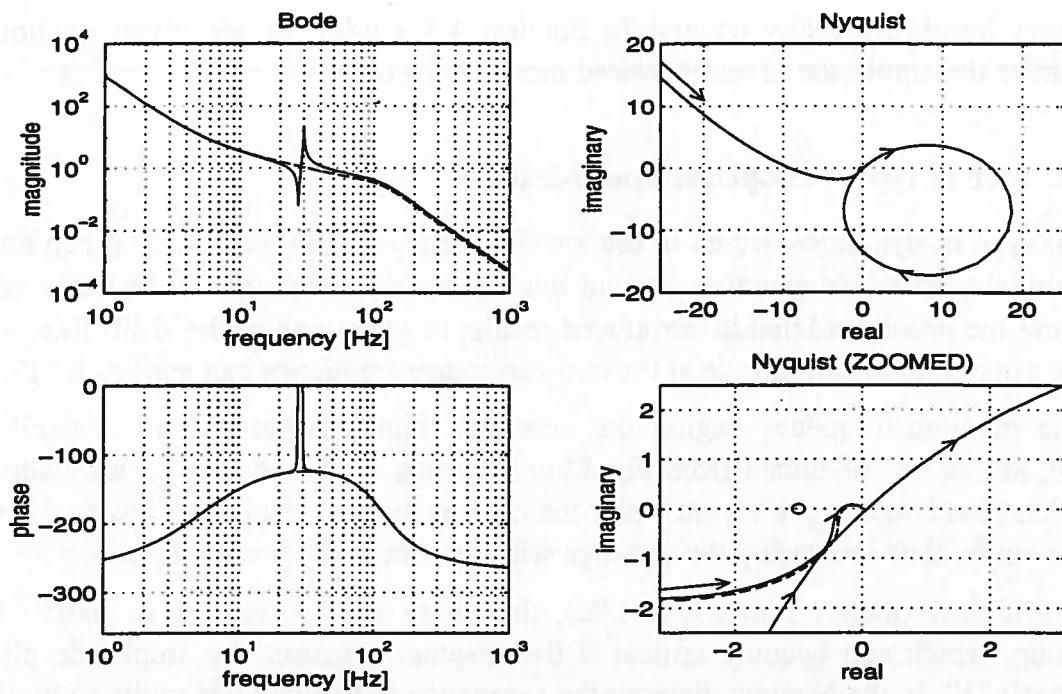


Figure 4.9b Open loop FRF of type “-2slope/zero/pole/-2slope” with medium-frequency resonance

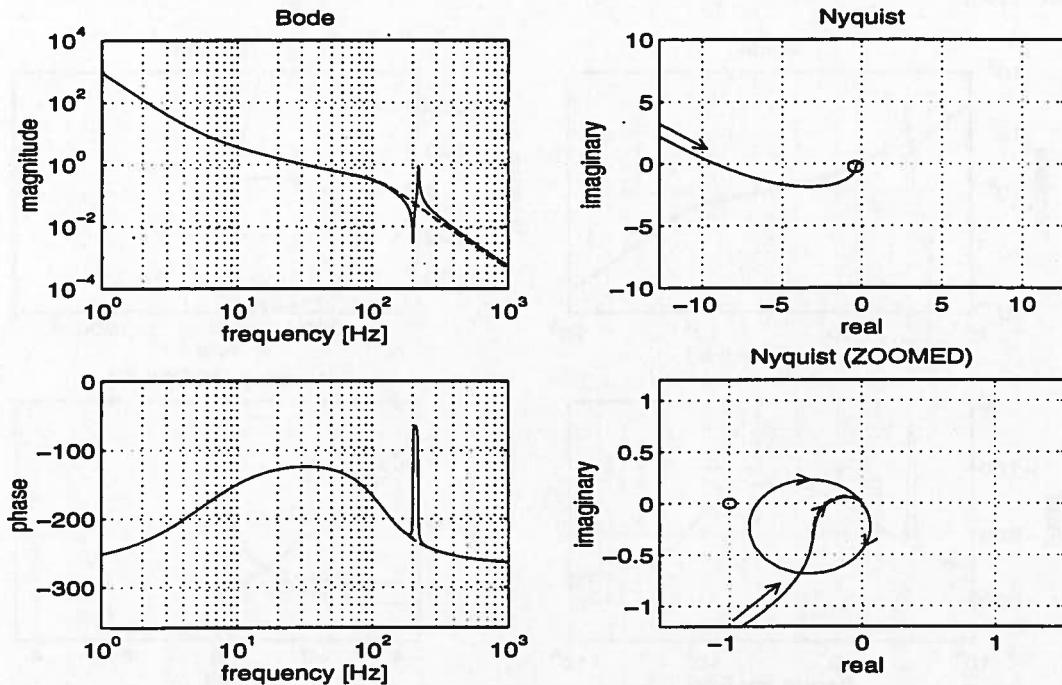


Figure 4.9c Open loop FRF of type “-2slope/zero/pole/-2slope” with high-frequency resonance

So far, the effect of the mass ratio α (equations (4.3) and (4.4)) has not been considered, and all plots have been generated with a small value of α , namely 0.1. In all cases a larger value of α , which indicates a larger importance of the additional mode, leads to a more destabilising effect. In the low-frequency and high-frequency regions a small value of α already leads to potential instability, therefore it is not

interesting to consider the influence of larger values. It is more important to understand whether the positive conclusion regarding the destabilising effect in the medium frequency range, drawn from Fig.4.9b, is still valid for larger values of α .

Figure 4.9d shows the resulting FRF for a mass ratio α equal to 5. As one can observe, the loop has become unstable. It is interesting to note that the zero/pole type of resonance phenomenon itself does not lead to this instability. It is the extra gain at high frequencies, which is identical to the mass ratio α , that causes the servo instability. This phenomenon can be explained by considering that the high-frequency 0-dB crossing has shifted towards a frequency range with no phase margin (alternatively, one could also state that the extra gain in the high-frequency range, where there is no phase lead, is responsible for the instability).

It depends very much on the transition frequency at which the phase changes from a lead to a lag, but this example shows that one has to be very cautious about the mass ratio α . In the discussion about design guidelines, Section 4.3, the issue of mass ratio will be raised again.

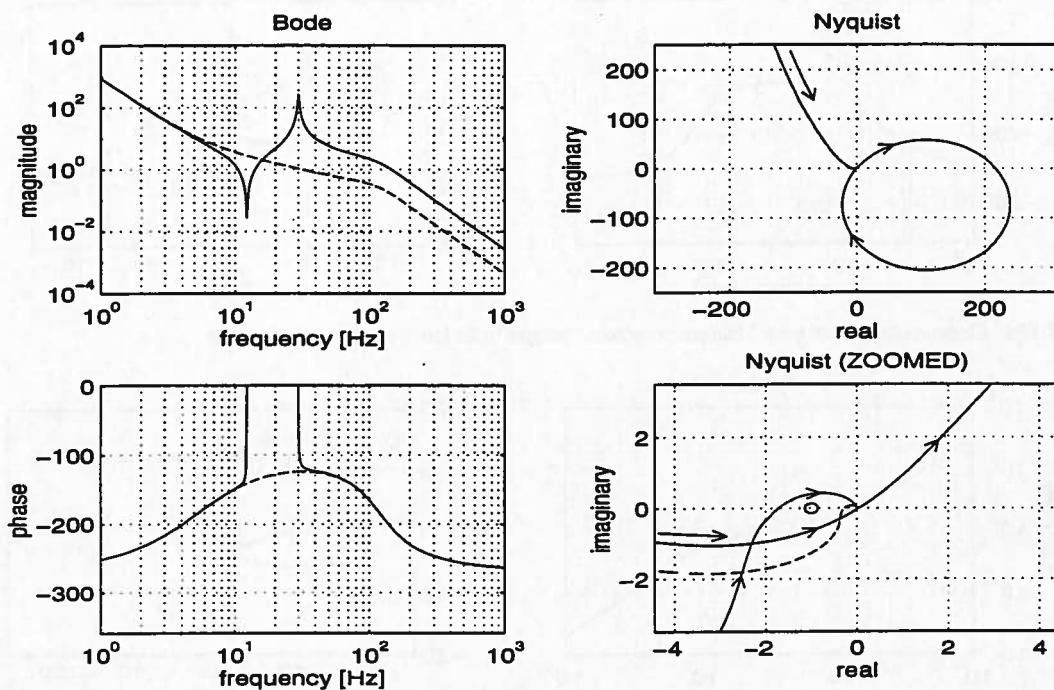


Figure 4.9d Open-loop FRF of type “-2slope/zero/pole/-2slope” with mass ratio $\alpha=5$

4.2.2. FRF of type “-2slope/pole/zero/-2slope”

The destabilising effect of the pole/zero combination is complementary, in terms of frequency range, to that of the previously discussed zero/pole combination.

At low frequencies (Fig.4.10a) the phase lag introduces no extra -180° crossing, and the resonance excursion, although lying on the left-hand side, has become small enough by the time the critical frequency region is reached.

However, in the medium-frequency region (Fig.4.10b) the pole/zero combination has a destabilising effect. In the Nyquist plot, the huge resonance means that the critical point (-1,0) can be passed on the wrong side.

In the high-frequency range (Fig.4.10c), there is no risk of instability. In the phase plot one can see that no extra -180° crossing is introduced, and the Nyquist plot shows that the resonance excursion happens in a safe part of the diagram.

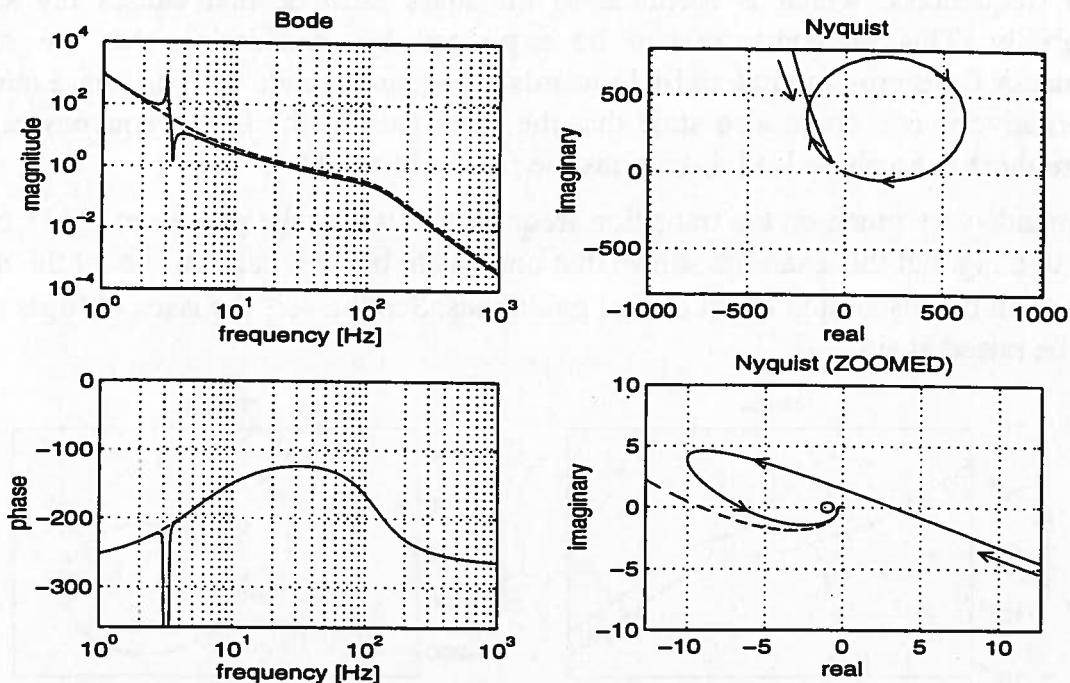


Figure 4.10a Open-loop FRF of type “-2slope/pole/zero/-2slope” with low-frequency resonance

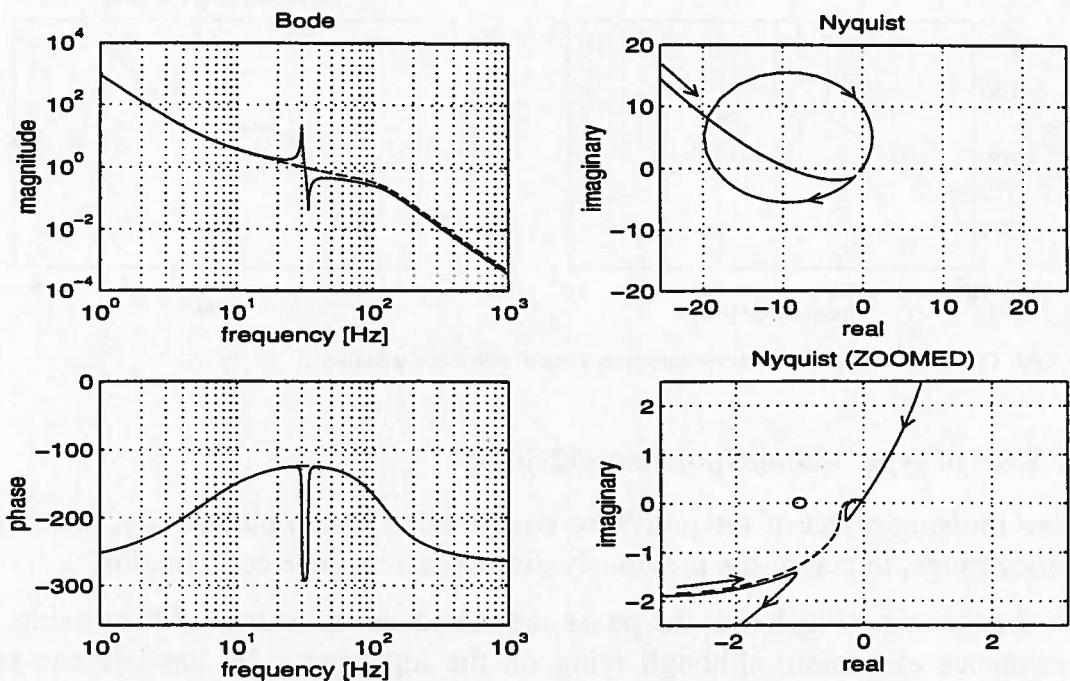


Figure 4.10b Open-loop FRF of type “-2slope/pole/zero/-2slope” with medium-frequency resonance

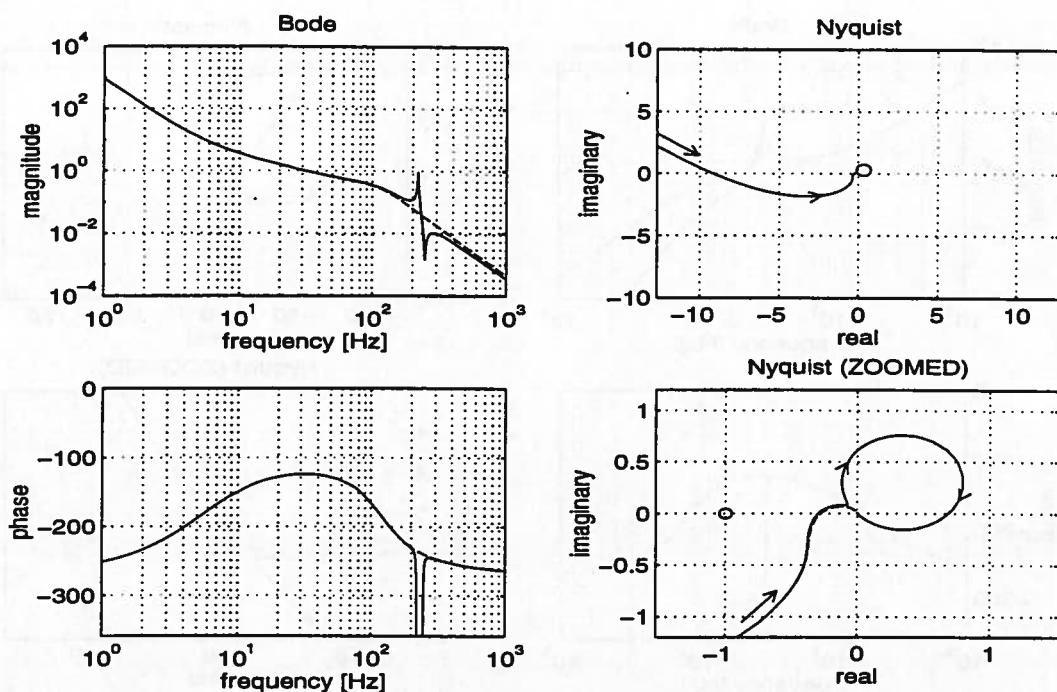


Figure 4.10c Open-loop FRF of type “-2slope/pole/zero/-2slope” with high-frequency resonance

4.2.3. FRF of type “-2slope/pole/-4slope”

This type of characteristic has a devastating effect on the stability of the control loop if the frequency of the resonance is not located high enough above the bandwidth frequency.

When the frequency is located in the low- or medium-frequency ranges, the loop will always be unstable, regardless of the amount of damping of the mode. Fig.4.11a shows what happens when the resonance is located in the medium-frequency range. As one can see in the Nyquist diagram, the critical point (-1,0) is circled on the wrong side.

If the resonance is located in the high frequency range (Fig. 4.11b), there is no reason for concern, as one can conclude from the Nyquist plot. If the resonance occurs at a frequency that is high enough compared with the frequency of the -180° crossing, then the corresponding excursion occurs in a safe region of the plot⁷.

⁷ In various publications it has been stated that for a stable operation of the loop, the amplitude at the resonance frequency must stay below “1”. This statement, however, is only true when the extra high-frequency phase lag (due to the second order filter, amplifier characteristics, or delay) is not considered, as one can conclude from the example.

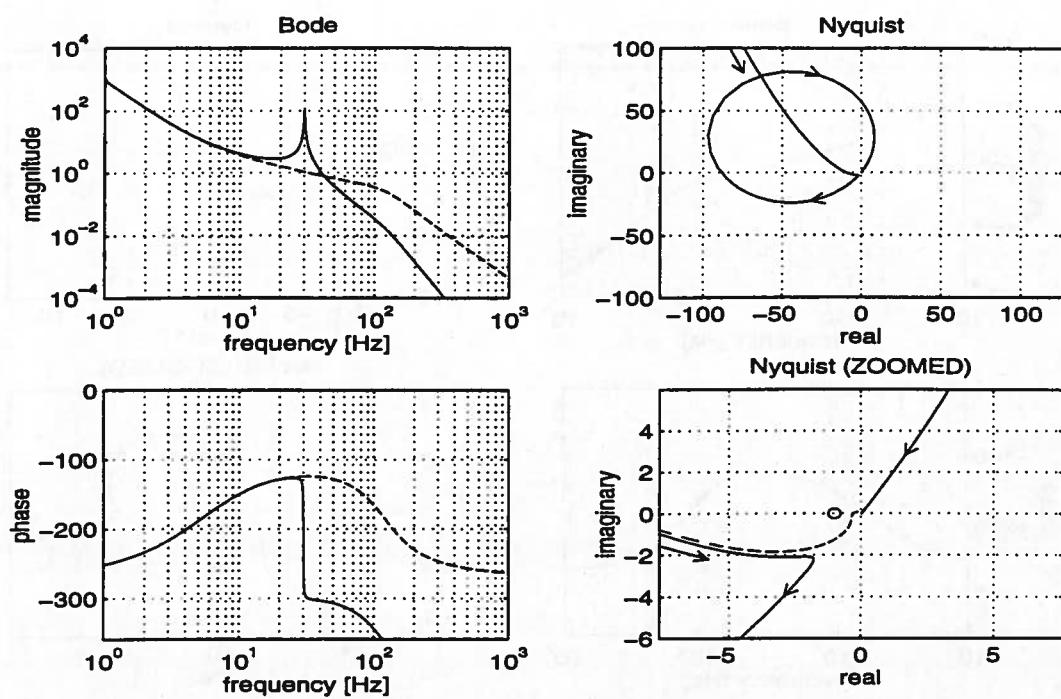


Figure 4.11a Open-loop FRF of type "-2slope/pole/-4slope" with medium-frequency resonance

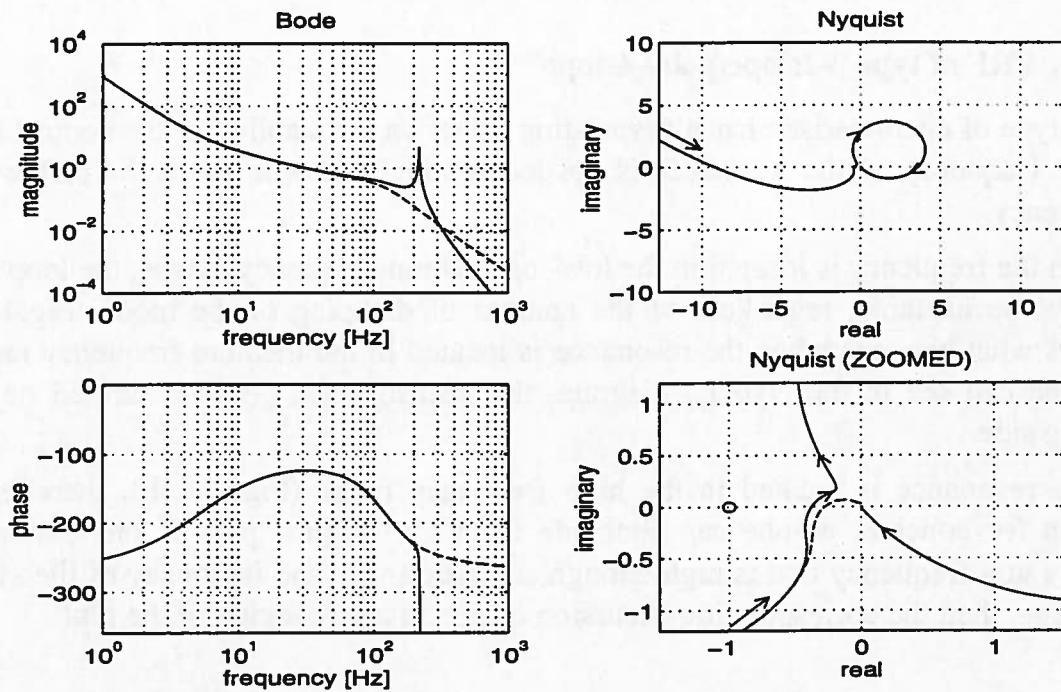


Figure 4.11b Open-loop FRF of type "-2slope/pole/-4slope" with high-frequency resonance

4.2.4. Resume

- A “-2 slope/zero/pole/-2 slope” characteristic leads to a phase lead, and is therefore potentially destabilising in the low-frequency and high-frequency regions. In the medium-frequency region it adds an extra phase lead to the already existing margin, which does not harm the stability
- A “-2 slope/pole/zero/-2 slope” combination has the reverse effect. It is potentially destabilising in the medium-frequency range and is harmless in the low- and high-frequency ranges.
- The “-2 slope/pole/-4 slope” behaviour always has a devastating effect on the stability of the loop if located in the low- or medium-frequency ranges. In the high-frequency range it can be dealt with.

Despite the fact that one specific controller has been used in the previous analyses, the results and conclusions have general validity and can be extrapolated to other controllers, provided that the three frequency regions are properly defined.

4.3. Design for Stability

The results of Section 4.2 can be linked to the three basic dynamic phenomena and practical design guidelines for the mechanical design can be given to minimise the destabilising effect on the control loop. These guidelines are based on the idea that the controlled system must satisfy certain disturbance rejection and bandwidth criteria. Whether this bandwidth can actually be achieved without introducing stability problems, depends very much on the mechanical design. In the final part of this section the issue of “design for stability” will be extended to more general dynamic effects, which are characterised by their modes of vibration.

4.3.1. Actuator Flexibility

Fig.4.12 shows the schematic representation of a system with a certain compliance between the motor and the load, and the corresponding modes, which were already derived in Chapter 3. Any FRF of this system can now be constructed by combining the contribution of each individual mode.

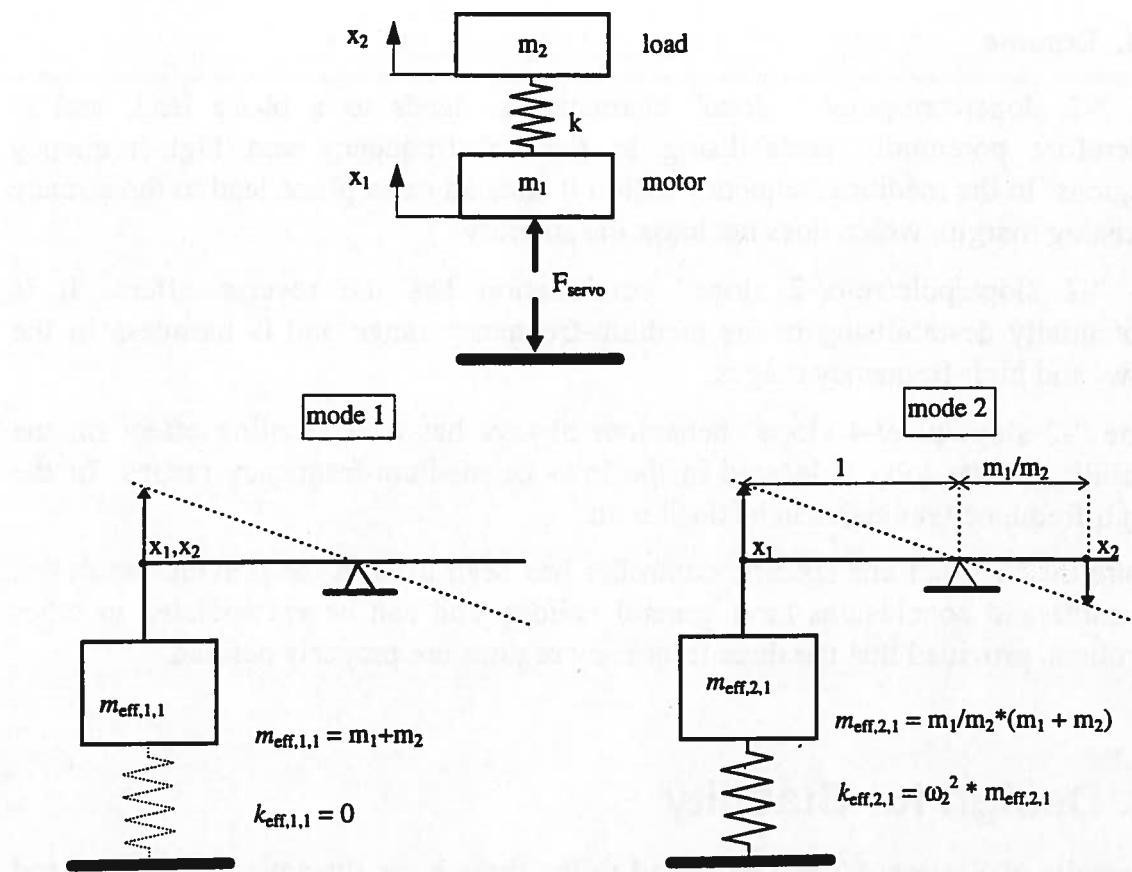


Figure 4.12 Servo System with Actuator Flexibility (a. schematic representation, b. modes)

Servo position measured at motor

Assuming that the feedback loop of the servo control system is based on measuring the position (or angle) of the motor, one has to investigate the following transfer function :

$$\left(\frac{x_1}{F_{\text{servo}}} \right) = \frac{1}{m_{\text{eff},11}s^2} + \frac{1}{m_{\text{eff},21}s^2 + \omega_2^2 m_{\text{eff},21}} \quad (4.9)$$

After substitution of the expressions for the effective modal masses from Fig.4.12, and some re-arrangement, this FRF can be expressed as :

$$\left(\frac{x_1}{F_{\text{servo}}} \right) = \frac{1}{(m_1 + m_2)s^2} + \frac{m_2/m_1}{(m_1 + m_2)*(s^2 + \omega_2^2)} \quad (4.10)$$

The mass ratio α (equations (4.3), (4.4)), which relates the “mass” of the additional modal contribution to the mass of the rigid-body motion, equals m_2/m_1 and is always positive. According to what is shown in Fig.4.5, the resulting FRF will therefore exhibit a -2slope/zero/pole/-2slope pattern (Fig.4.13).

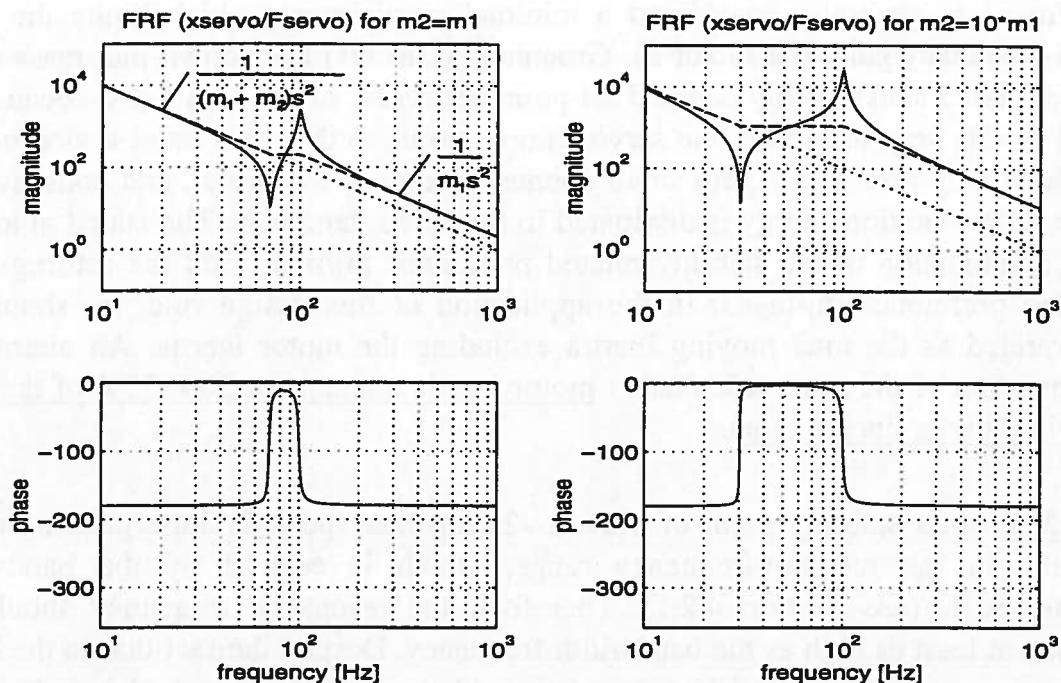


Figure 4.13 Mechanical FRF of a system with actuator flexibility and position measurement at motor for two values of the mass ratio m_2/m_1

The asymptotes at low and high frequencies are :

$$\left(\frac{x_1}{F_{\text{servo}}} \right)_{s \rightarrow 0} = \frac{1}{(m_1 + m_2)s^2} \quad (4.11)$$

$$\left(\frac{x_1}{F_{\text{servo}}} \right)_{s \rightarrow \infty} = \frac{1}{(m_1 + m_2)s^2} + \frac{1}{m_1/m_2(m_1 + m_2)s^2} = \frac{1}{m_1 s^2} \quad (4.12)$$

which correspond to the engineering feeling that at very low frequencies the two masses move as one single mass, whereas at very high frequencies the mass m_2 of the load is completely decoupled such that the servo system only “feels and sees” the motion of the motor mass m_1 . With respect to the design of such a configuration, the following guidelines can be given, which are a combination of stability and set-point response criteria :

Actuator flexibility with measurement at motor

1. motor inertia = (1..3) * load inertia
2. resonance frequency $f_e \geq$ bandwidth f_b

Ad 1. In the discussion of the destabilising effect of this type of FRF (Section 4.2.1) the effect of the mass ratio was already mentioned. The extra gain at high frequencies due to the decoupling of the load mass m_2 (see Fig.4.13), can have a negative effect on the stability. From a stability point of view, it is therefore advisable to minimise the extra gain by choosing the ratio m_1/m_2 as high as possible

$(m_1/m_2=1)$ is generally considered a minimal requirement, which limits the extra high-frequency gain to a factor 2). Groenhuis [GRO91] has shown that mass ratios larger than 3 lead to badly damped set-point behaviour of the load mass, because the load profits very little from the servo damping (due to the mass ratio, a vibration of the load only results in a very small counter motion of the motor, and consequently only little vibration energy is dissipated in the servo damping). The stated guideline is a combination of the stability related preference $m_1/m_2>1$ and the settling-time-related preference $m_1/m_2<3$. In the application of this design rule, m_2 should be interpreted as the total moving inertia excluding the motor inertia. An alternative formulation of the same rule reads : motor inertia amounts to 50%-75% of the total moving inertia (incl. motor).

Ad 2. From a stability point of view a -2 slope/zero/pole/-2 slope pattern is best located in the medium-frequency range, which is centred by the bandwidth frequency f_b (see Section 4.2.1). Therefore, the resonance frequency should be chosen at least as high as the bandwidth frequency. Despite the fact that in the high-frequency range where the phase has dropped below -180° , a zero/pole combination can potentially lead to instability, it is advisable to aim at a high mechanical resonance frequency. First, because generally this gives the best set-point response, and second, because one can still obtain a stable control loop if the resonance frequency is high enough (and therefore the amplitude of the open loop at the resonance is sufficiently low). As an alternative, one could move the resonance from the high-frequency range to the medium-frequency range by increasing the bandwidth frequency (if possible) or simply by lowering the mechanical-resonance frequency (lowering a mechanical-resonance frequency is generally much easier than increasing it).

Servo position measured at load

In some applications the position of the load is used in the feedback loop, and therefore one has to consider the transfer function (x_2/F_{servo}) , which again can be derived very easily from the modal representation (Fig.4.12). As the location of the servo force F_{servo} has not changed, one can basically maintain expression (4.9) that was derived earlier for (x_1/F_{servo}) and multiply each modal contribution with the corresponding scaling factor in order to obtain (x_2/F_{servo}) :

$$\left(\frac{x_2}{F_{\text{servo}}} \right) = 1 * \frac{1}{m_{\text{eff},11}s^2} + \left(-\frac{m_1}{m_2} \right) * \frac{1}{m_{\text{eff},21}s^2 + \omega_2^2 m_{\text{eff},21}} \quad (4.13)$$

$$\left(\frac{x_2}{F_{\text{servo}}} \right) = \frac{1}{(m_1 + m_2)s^2} + \frac{-1}{(m_1 + m_2)(s^2 + \omega_2^2)} \quad (4.14)$$

Equation (4.14) reveals that when the servo position is measured at the load, the mass ratio α equals -1, and thus the FRF will be of the type “-2slope/pole/-4slope”. At frequencies below the resonance frequency, the system behaves practically as one rigid body. However above the resonance frequency, the load cannot “follow” the motion of the motor anymore, and is decoupled. Fig.4.14 shows the Bode diagram of this transfer function multiplied by a proportional gain k_p .⁸

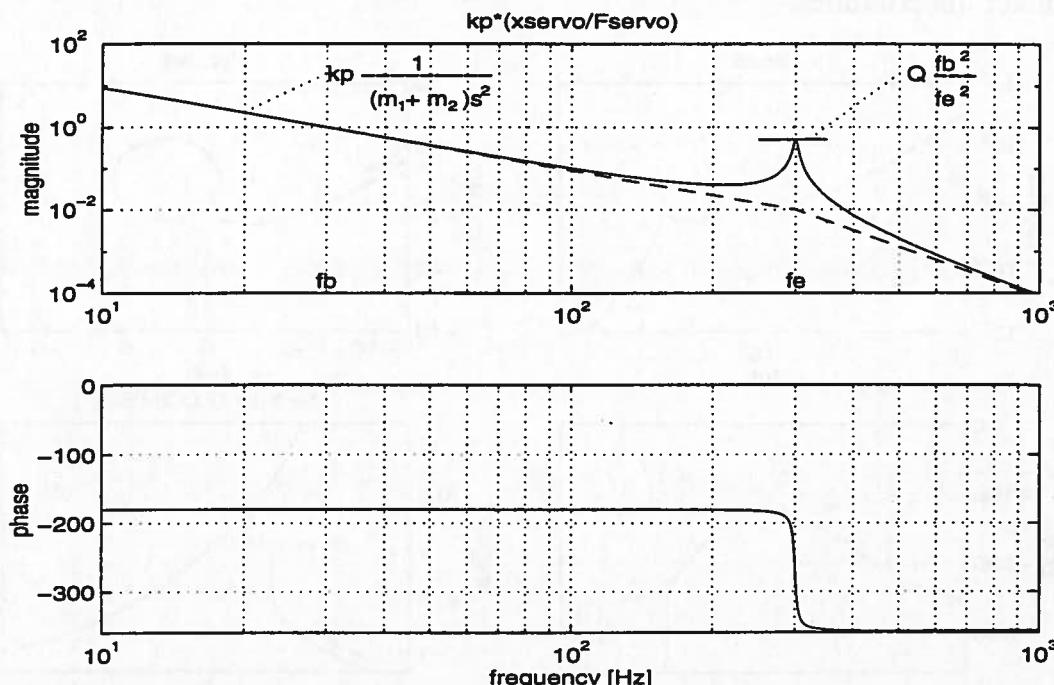


Figure 4.14 FRF $k_p \cdot (x_{\text{servo}}/F_{\text{servo}})$ of a system with actuator flexibility and position measurement at the load

For this type of configuration, the following design guideline can be given :

Actuator flexibility with measurement at load

resonance frequency $f_e > (5..10) \cdot$ bandwidth f_b

which can be derived from the fact that a stable control loop can only be achieved when at least one of the following two conditions is met :

1. sufficient phase lag at the resonance frequency such that the resonance excursion in the Nyquist diagram lies in a safe region
2. sufficient amplitude margin at the resonance frequency

In Section 4.2.3 it was already demonstrated that despite a resonance peak that crosses the 0-dB line, a control loop can still be stable if sufficient phase lag is present at the resonance frequency, such that the resonance excursion lies in a safe region of the Nyquist diagram. The conflicting requirements of a phase lead at the

⁸ The multiplication with k_p has been done in view of the discussion about design guidelines.

bandwidth frequency f_b and a phase lag at the resonance frequency f_e can only be met if the ratio f_e/f_b is large enough (somewhere between f_b and f_e a -180° crossing of the phase must be realised). Figure 4.15 shows the open-loop FRF for a ratio $f_e/f_b=5$, which can be considered a minimum value in order to achieve a stable system. Generally, it is advisable to aim at a higher ratio, because the positive results on the basis of $f_e/f_b=5$ can only be achieved for a well-tuned system with very few parameter uncertainties.

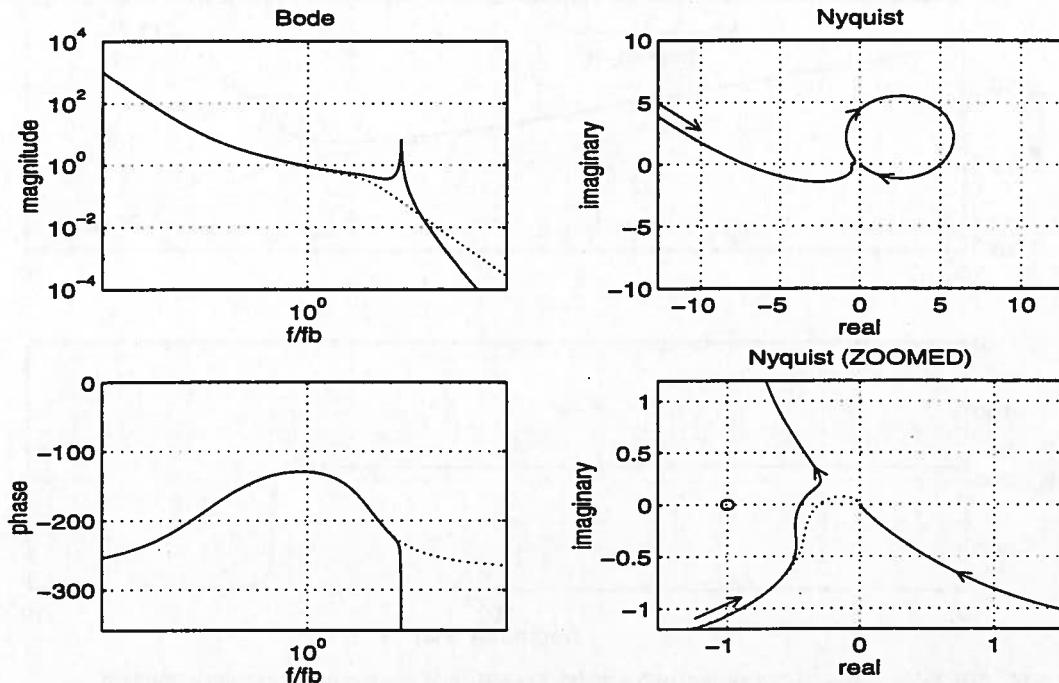


Figure 4.15 Open loop FRF of a system with actuator flexibility and position measurement at the load ($f_e/f_b=5$)

Alternatively, a stable control loop can also be obtained by satisfying the second condition, which can be translated into a design guideline on the basis of the following engineering reasoning: By only considering the proportional gain of the controller, and defining the bandwidth frequency f_b as the frequency of the 0-dB crossing, the amplitude at the resonance frequency f_e can be approximated by (see Fig.4.14) :

$$|H(f_e)| \approx Q * (f_b/f_e)^2 \quad (4.15)$$

in which Q represents the amplification factor that is related to the relative damping via $Q=1/(2\beta)$. Values for the relative damping β are generally in the range of 1%-2%, so Q ranges from 25-50. In combination with an amplitude margin of 6dB, equation (4.15) leads to the following specification for the resonance frequency relative to the required bandwidth :

$$|H(f_e)| \approx 50 * (f_b/f_e)^2 < 0.5 \implies f_e/f_b > 10 \quad (4.16)$$

4.3.2. Guiding System Flexibility

The modes of a servo actuator with limited guiding stiffness (Fig.4.16) were already discussed in Chapter 3. Here, the influence of the rocking mode on the FRF of such an actuator system will be analysed.

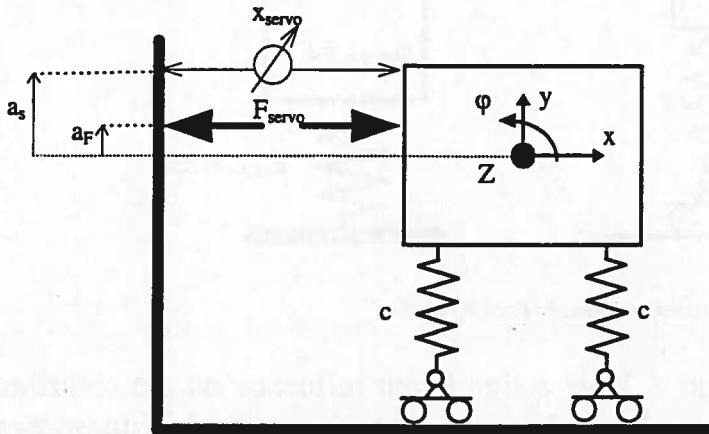


Figure 4.16 2D Rigid-Body Model of Actuator with Flexibility of the Guiding System

Assuming that the servo force F_{servo} acts at a certain distance a_F with respect to the centre of gravity, indicated by the point Z , and the servo position is measured at a distance a_s with respect to the centre of gravity, then the overall open-loop transfer function will certainly suffer from the effect of the rocking mode. Due to the assumed symmetry of the system, the Y motion is completely decoupled from the X and ϕ motions of the system, and can therefore be omitted in this analysis.

Considering the two relevant modes (Fig.4.17a, Fig.4.17b), the resulting transfer function ($x_{\text{servo}}/F_{\text{servo}}$) can be constructed from the contributions of the individual modes. With the knowledge of the mode-shapes it is easy to set up the equation for this transfer function, in which the first part accounts for the desired X motion, and the second part represents the effect of the rocking mode :

$$\left(\frac{x_{\text{servo}}}{F_{\text{servo}}} \right) = \frac{1}{m s^2} + \frac{a_s a_F}{J s^2 + 2cb^2} \quad (4.17)$$

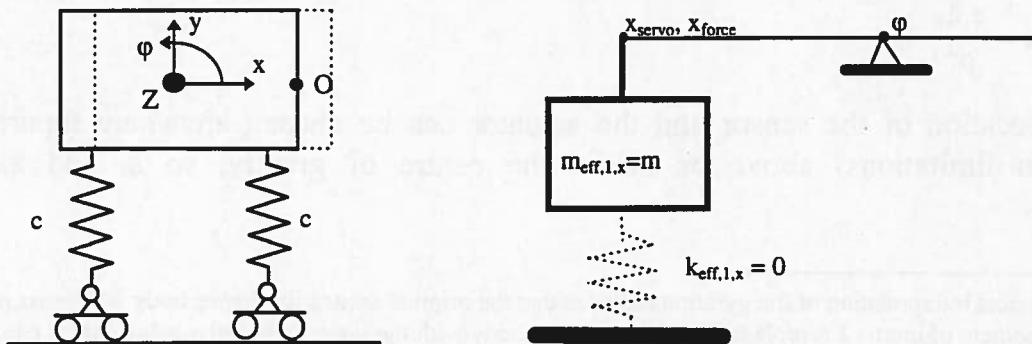


Figure 4.17a Graphical Representation of Desired X-Motion

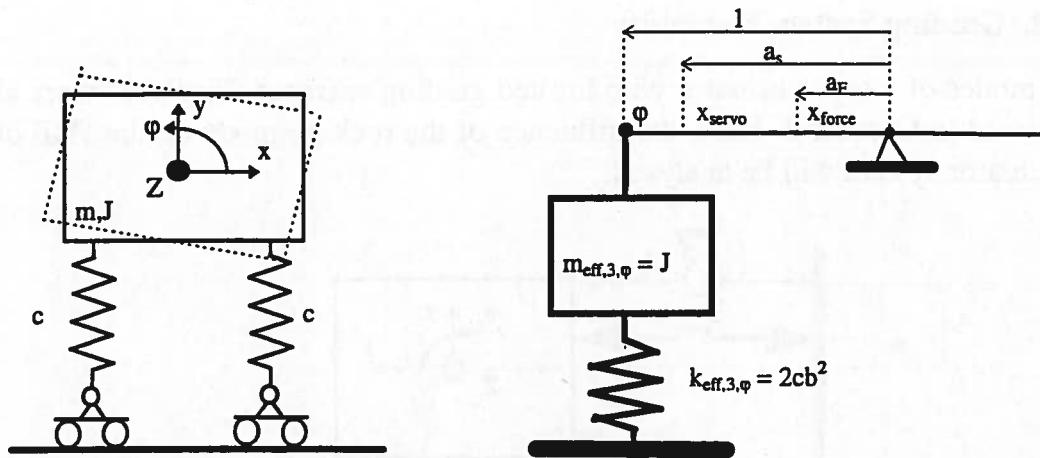


Figure 4.17b Graphical Representation of Parasitic Rocking Mode

The two distances a_F and a_s have a significant influence on the contribution of the rocking mode on the open loop characteristics (the product of these two distances and not the individual values are important, so exchanging the position of the sensor and the force does not affect the response).

Before constructing the overall transfer function, it is useful to introduce the so-called gyration radius ρ , which links the mass moment of inertia J to the mass m of the system in the following way⁹:

$$J = m * \rho^2 \quad (4.18)$$

Using the gyration radius in the equation of the transfer function $x_{\text{servo}}/F_{\text{servo}}$ leads to:

$$\left(\frac{x_{\text{servo}}}{F_{\text{servo}}} \right) = \frac{1}{ms^2} + \frac{a_s a_F}{2cb^2 + mp^2 s^2} \quad (4.19)$$

At high frequencies, this formula can be approximated by :

$$\left(\frac{x_{\text{servo}}}{F_{\text{servo}}} \right)_{s \rightarrow \infty} = \frac{1}{ms^2} + \frac{a_s a_F}{mp^2 s^2} = \frac{1}{ms^2} \left(1 + \frac{a_s a_F}{\rho^2} \right) \quad (4.20)$$

Equation (4.20) reveals that for this type of system, the mass ratio α equals :

$$\alpha = \frac{a_s a_F}{\rho^2} \quad (4.21)$$

The location of the sensor and the actuator can be chosen anywhere (apart from design limitations) above or below the centre of gravity, so a_s and a_F , and

⁹ A physical interpretation of the gyration radius is that the original arbitrarily shaped body with mass m and mass moment of inertia J is replaced by a ring-shaped body with the same mass and a radius that is equal to the gyration radius of the original system. The mass properties of this new body are identical to those of the original system.

consequently α , can become any positive or negative value, and the resulting FRF can have any of the characteristics shown in Fig.4.5.

As long as the actuating force and the sensor are located on the same side of the centre of gravity, α will always be positive, and the overall FRF will display “-2slope/zero/pole/-2slope” behaviour. Where the force and sensor are located at opposite sides of the centre of gravity, one of the other three characteristic shapes will be found, depending on the exact values of a_s and a_F . If the force and the sensor are both located outside the gyration radius and on opposite sides of the centre of gravity ($|a_s|>p$, $|a_F|>p$ and $\text{sign}(a_s a_F)=-1$), equation (4.21) will even yield a mass ratio $\alpha<-1$. According to the overview (Fig.4.5), such a mass ratio leads to a “-2slope/pole/ -2slope” FRF, which displays a non-minimum phase behaviour (see Section 4.1). As one can see, practically any FRF can be generated by such a dynamic system in which the rocking mode due to guiding flexibility plays an important role.

The following general design rules apply for such a system¹⁰ :

Guiding system flexibility

- 1. Driving force at COG (BEST PRACTICE)**
- 2. Locate sensor at COG (Second Best)**
- 3. If none of the above can be achieved one should aim at locating sensor and driving force as close as possible to the COG. Furthermore, it is generally better if the location of the sensor and that of the driving force are at the same side of the COG.**

Ad 1. The best way to avoid any dynamic problems is to drive the system at its COG. By doing so the rocking mode is not excited, which is not only favourable from a stability point of view, but also results in good set-point behaviour.

Basically this guideline is a special case of a more general rule that states that “driving forces of a rigid actuator should be applied in such a way that the actuator performs the desired motion even in the absence of any guiding system”. A closer look at this rule reveals that its application automatically guarantees that the driving force is located at the node of any undesired mode which is related to the limited stiffness of the guiding system. As an illustration of this more general guideline, Fig.4.18 shows an actuator (mass m and mass moment of inertia J at centre of gravity) with leaf-spring guiding and its equivalent representation with a ball bearing. The actuator is free to rotate about point O, whereas the other two DOF are restrained by the guiding system.

¹⁰ The choice of sensor location and driving force location will also be discussed in a more general manner in Section 4.3.4.

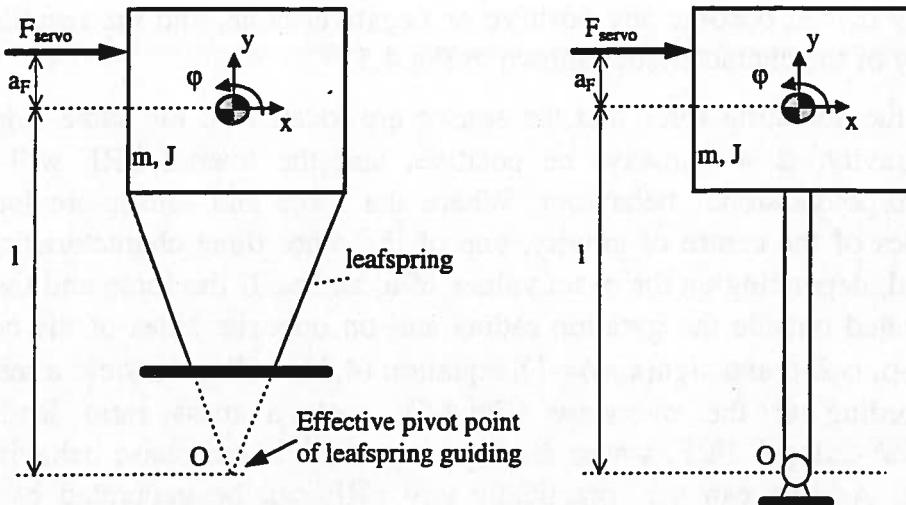


Figure 4.18 Actuator with leaf-spring guiding and equivalent representation with ball-bearing

Due to the constraints imposed by the guiding system, the static relation between the x-motion and the φ -motion of the COG is given by :

$$\varphi = -\frac{x}{l} \quad (4.22)$$

Therefore, according to the design rule, the driving force F must be applied at such a distance a_F above the COG that the accelerations satisfy the following relation :

$$\ddot{\varphi} = -\frac{\ddot{x}}{l} \quad (4.23)$$

The accelerations that follow from the driving force F are :

$$\ddot{\varphi} = -\frac{F a_F}{J} \quad \text{and} \quad \ddot{x} = \frac{F}{m} \quad (4.24)$$

and therefore in order to satisfy equation (4.23) the distance a_F must be chosen according to :

$$a_F = \frac{J}{m l} \quad (4.25)$$

As already mentioned, the point at a distance a_F above the COG also has a special meaning when investigating the modes of the system that result from a limited stiffness of the support. In the case of a translational stiffness c in the X and Y directions of the (effective) pivot point, two modes will exist next to the desired rigid-body rotation about the pivot point. Due to the assumed symmetry of the system, one of these modes consists of a motion in the Y direction only, and because there is no coupling with the X and φ motions it will not be considered here. The other mode is a combined X and φ motion which results in a rocking mode with a node that is located exactly at a distance a_F (see equation (4.25)) above the COG, as can be derived from the corresponding eigenvalue problem :

$$\begin{bmatrix} m & 0 \\ 0 & J \end{bmatrix} \begin{bmatrix} \ddot{x} \\ \ddot{\phi} \end{bmatrix} + \begin{bmatrix} c & cl \\ cl & cl^2 \end{bmatrix} \begin{bmatrix} x \\ \phi \end{bmatrix} = 0 \quad (4.26)$$

which can be transformed into :

$$\begin{bmatrix} c/m & cl/m \\ cl/m & cl^2/J \end{bmatrix} \phi_i = \omega_i^2 \phi_i \quad (4.27)$$

Solving equation (4.27) yields two natural frequencies and corresponding modes, which for convenience have been scaled such that the ϕ component is equal to 1 :

$$\phi_1 = \begin{vmatrix} -1 \\ 1 \end{vmatrix} \quad \omega_1^2 = 0 \quad (4.28a)$$

$$\phi_2 = \begin{vmatrix} +J/ml \\ 1 \end{vmatrix} \quad \omega_2^2 = \frac{c}{m} + \frac{cl^2}{J} \quad (4.28b)$$

Using the relation for the distance a_F according to (4.25) the mode-shape vector ϕ_2 (4.28b) can also be written as :

$$\phi_2 = \begin{vmatrix} a_F \\ 1 \end{vmatrix} \quad (4.29)$$

which indeed corresponds to a rotation of the actuator body about a point that is located at a distance a_F above the COG.

Ad 2. When it is impossible to apply the forces at the correct location, one can eliminate the destabilising effect of the rocking mode by locating the sensor at the height of the COG. As this is a stationary point, the resonance will not be present in the FRF.

Ad 3. The combination of functions of such a guiding system, namely unlimited motion in the X direction and suppressed motion in the other directions, will limit the realisable stiffness of the support and thus the resonance frequency of the rocking mode. Consequently, the resonance frequency will often fall within the medium-frequency range. In this frequency range a -2 slope/zero/pole/-2 slope characteristic is favourable, so the location of the driving force and that of the sensor should be chosen at the same side of the COG.

When one has to settle for a configuration in which neither the sensor nor the driving force can be located at the COG, it is useful to have some indication of the tolerable offsets. If one aims at a phase margin of about 45° , a deviation of this value by 5° due to a resonance seems an acceptable engineering compromise. Assuming a relative damping $\beta=1\%$, such a phase deviation already occurs for a mass ratio $\alpha=0.002$. With equation (4.21), and choosing $|a_s|=|a_F|$, this value of α leads to a tolerable offset of about $0.02*L$ in the case of a square actuator (see fig.4.16) with dimension $L*L$.

4.3.3. Limited Mass and Stiffness of Stationary Machine Part

Figure 4.19 shows a simple model of a translational direct drive motor on a frame with limited mass and stiffness, and in which the control system operates on the measured position $x_{\text{servo}} = x_2 - x_1$.

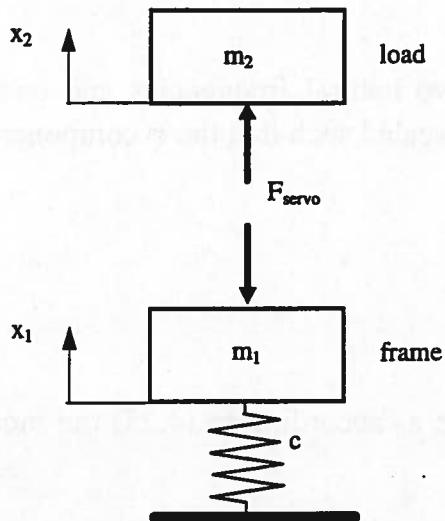


Figure 4.19 Model of a servo system including frame dynamics

The two modes of this system are the 0 Hz rigid-body motion of mass m_2 and the vibration of the frame mass m_1 on the frame stiffness c . The influence of the frame dynamics on the stability of the control loop can be judged on the basis of the transfer function ($x_{\text{servo}}/F_{\text{servo}}$), which reads :

$$\left(\frac{x_{\text{servo}}}{F_{\text{servo}}} \right) = \left(\frac{x_2 - x_1}{F_{\text{servo}}} \right) = \frac{1}{m_2 s^2} + \frac{1}{m_1 s^2 + c} = \frac{1}{m_2 s^2} + \frac{m_2/m_1}{m_2 s^2 + c(m_2/m_1)} \quad (4.30)$$

The mass ratio α equals m_2/m_1 and is therefore always positive. Consequently, the resulting transfer function is of type “-2slope/zero/pole/-2slope”. The asymptotes at low and high frequencies are :

$$\left(\frac{x_{\text{servo}}}{F_{\text{servo}}} \right)_{s \rightarrow 0} = \frac{1}{m_2 s^2} \quad (4.31)$$

$$\left(\frac{x_{\text{servo}}}{F_{\text{servo}}} \right)_{s \rightarrow \infty} = \frac{1}{m_1 s^2} + \frac{1}{m_2 s^2} = \frac{m_1 + m_2}{m_1 m_2 s^2} \quad (4.32)$$

The following design rules can be given :

Frame Motion

1. frame inertia > load inertia (measurement between load and frame)
2. frame inertia > 50 * load inertia (measurement between load and additional flexible structure attached to the frame)

Ad 1. For this type of FRF it was previously stated that, from a stability point of view, the extra gain above the resonance frequency should be as small as possible (an upper limit of a factor 2 is considered a reasonable engineering requirement). From equations (4.31) and (4.32) it follows that the extra gain equals $(m_1+m_2)/m_1$. Therefore, in order to obtain an extra gain that is smaller than 2, the frame mass m_1 should be at least equal to the load mass m_2 , and preferably higher. Under normal conditions it should not be difficult to design the machine such that this requirement is fulfilled¹¹.

Ad 2. The effect of the reaction forces on the frame can put even higher demands on the frame mass when the position measurement does not take place directly between the load and the frame, but rather between the load and some extra component that is flexibly connected to the frame. In such a situation, which is described in more detail in the concept analysis of the case study (Section 6.3), it is favourable to have a frame mass larger than Q times the load mass, with Q being the resonance amplification of the additional component. With Q values ranging from 25-50 (relative damping between 1%-2%) this leads to the mentioned design rule. The consequence of not fulfilling this second, more stringent, design rule, which becomes important as soon as there is potential flexibility between the frame and the location of the sensor that measures the position of the load, can be quite dramatic. In such a case it will be necessary to design, manufacture, and assemble all parts that are attached to the stationary machine frame with the utmost care in order to avoid dynamics related problems.

It is important to note that the guidelines for this type of configuration are based on stability criteria only. Because frame vibrations can have a significant influence on the set-point response, this aspect should also be evaluated (see Appendix A).

Final Remarks

The presented guidelines are intended to guide the design of the mechanical system in the right direction and in the majority of cases they will be very helpful. However, they should not be interpreted as rigid rules which must be obeyed, because there are always exceptional situations in which a different solution is better.

Despite the fact that on the basis of combined mechanics/control considerations, guidelines have been derived for the mechanical design, there must always be a close co-operation between the mechanical designer and the control engineer. Only with the input of both disciplines can the optimum mechatronic balance between mechanical and control efforts be realised.

¹¹ Note that the same type of characteristic is also observed when the stationary part of a motor (for example the stator of a rotating motor) is insufficiently stiffly connected to the (heavy) frame. In such a situation it is not evident that the design rules regarding the mass ratio between moving and stationary mass are satisfied.

4.3.4. General Guidelines

This chapter focuses on approaches that minimise or even eliminate the destabilising effect of arbitrary modes.

The amount of contribution of a certain mode (Fig.4.20) to the open-loop response (x_{servo}/F_{servo}) and its interaction with the desired motion is determined by the modal mass and stiffness, but also by the location of the driving force (resulting in a moment on the modal lever) and the location of the response DOF.

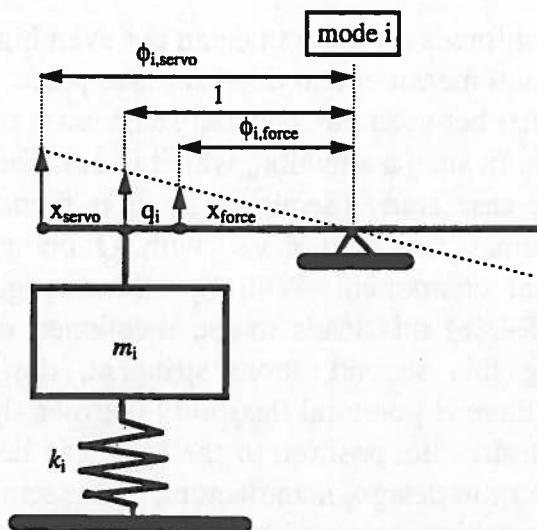


Figure 4.20 Graphical representation of mode i

This observation leads to the idea that an undesired contribution of a mode to the response can be eliminated by one of the following two approaches :

1. Mode should not be excited, which implies that the total moment acting on the modal lever should be equal to zero :
 - Locate driving force at a node of the mode.
 - Modify mode-shape such that the location of the driving force becomes a node of the mode.
 - Apply additional excitation force(s) such that the overall moment acting on the modal lever is equal to zero.
2. Output of response DOF should be zero :
 - Shift response DOF (sensor) towards node of mode
 - Modify structural system such that the sensor location becomes a node of the mode
 - Add additional sensor(s) and combine the outputs such that the contribution of the mode to the new response signal is equal to zero. A very nice example of this approach is described in [STE96].

It can be shown [SET89] that these two approaches are closely related to the terms “(un)controllability” and “(un)observability” that are frequently used in modern control theory. Uncontrollability of a mode corresponds to the first approach (driving force has no effect on the mode), whereas unobservability is obtained in the second approach (sensor is located at stationary point of a mode). From control theory it is known that only those modes can be influenced or modified by the feedback loop which are observable and controllable. If such a modification is not required and the modes are not excited by some other mechanism, it can be very effective to use uncontrollability and unobservability in a positive way in order to eliminate unwanted resonances that could endanger stability.

Some of the approaches that can be used to avoid the destabilising effect of machine dynamics have been applied successfully in various projects. As an illustration, a simulation study [JON93] of an optical pickup unit from a Compact Disc drive (Fig.4.21 and Fig.4.22) will be shown.

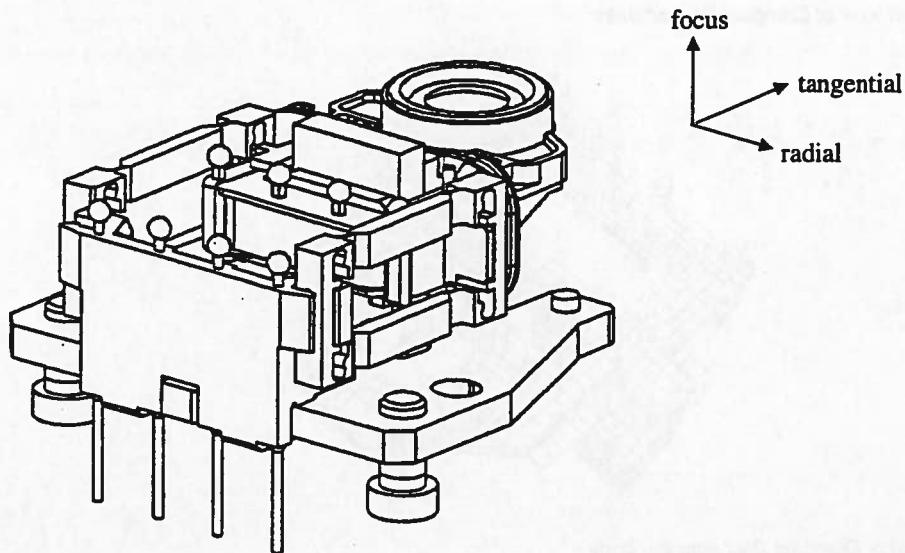


Figure 4.21 Compact Disc actuator

The 2D actuator is controlled in vertical (focus) direction in order to focus the laser spot on the disc, and in radial (tracking) direction in order to keep the laser spot on the track. The actuator is supported by four suspension bars, which are stiff in axial direction and flexible in the perpendicular directions. It is driven by one focus coil and a set of two radial coils, that move within a magnetic field.

Typical bandwidth of the radial and focus controls is 1-3kHz, which implies that internal resonances of the actuator, which start at about 10kHz can have a destabilising effect on the control loop. To understand and improve its internal dynamics, in particular in the radial direction, a FE model of the actuator body has been created (Fig.4.23). As the stiffness of the suspension bars is negligible at 10kHz, the actuator body is analysed as an unsupported system.

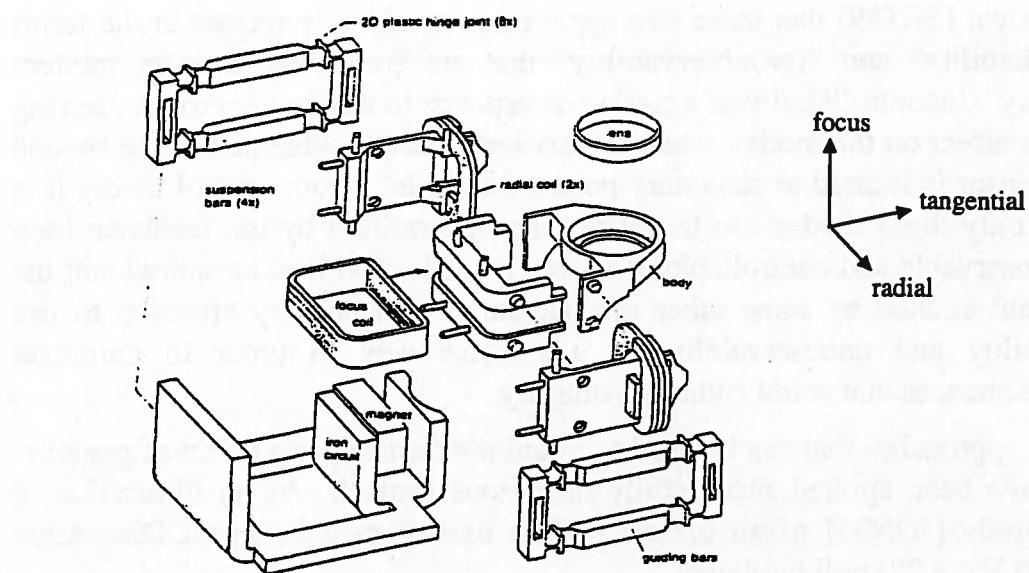


Figure 4.22 Exploded view of Compact Disc actuator

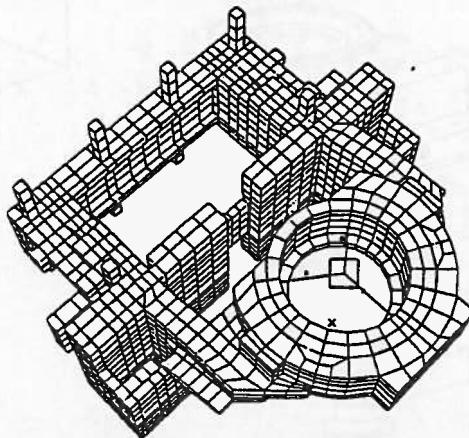


Figure 4.22 FE-model of Compact Disc actuator body

The simulated radial FRF ($x_{\text{lens}}/F_{\text{coil}}$) of the actuator is shown in Fig.4.23. There are two important modes¹², namely at 9.4kHz and at about 12kHz. As one can derive from the amplitude and phase relation, the lens is completely decoupled above 12kHz. Aiming at a bandwidth frequency of about 2kHz, one can assume that both resonances are located in the frequency range in which the phase of the entire loop has already dropped below -180° (see Fig. 4.8). Consequently, a phase lead, as introduced by the zero/pole combination at 9.4kHz, is a serious threat for stability. The decoupling at 12kHz on the other hand does not pose any problem, as the frequency is high enough compared with the bandwidth frequency. The Nyquist diagram exposes the problem from a different angle. It shows that the two resonances display opposite phase excursions, which makes it practically impossible to avoid instability by applying a phase action.

¹² The contribution of the first resonance peak is considered negligible.

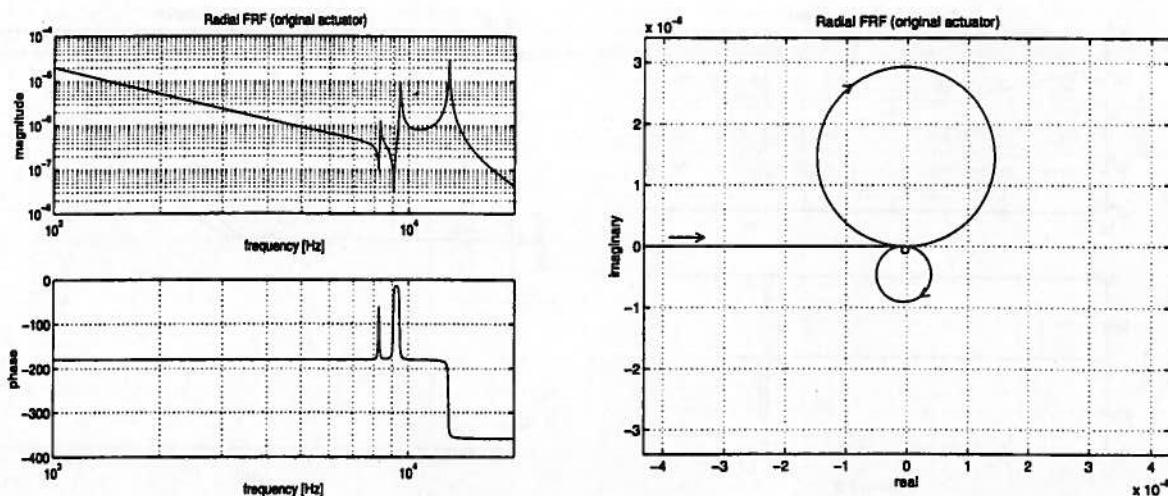


Figure 4.23 Radial FRF $X_{\text{lens}}/F_{\text{radial coil}}$ of original actuator (left: Bode diagram, right: Nyquist diagram)

An investigation of the mode-shape that corresponds to the 9.4kHz resonance (Fig.4.24), reveals that the radial displacements of the excitation DOF (= location of radial coils) and the relevant response DOF (= centre of lens) are indeed in phase. In previous chapters it was shown that this must be the case in order to find a zero/pole combination in the FRF.

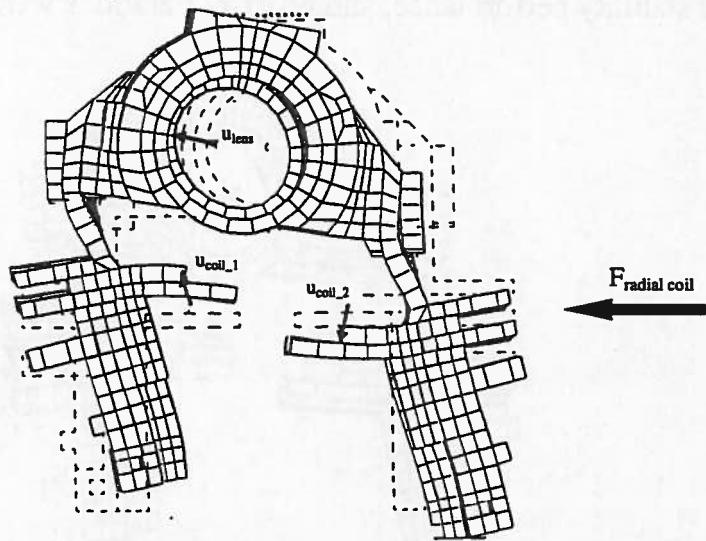


Figure 4.24 Mode shape at 9.4 kHz (original actuator)

After some iterations, which were first focused on increasing the frequency of that mode, attempts were made to reduce the excitation of that mode via structural modifications. In one of the simulations, some material has been removed at the lens-side of the bars that clearly deform most, which leads to a mode with a slightly lower frequency, 8.8kHz, but with a more favourable¹³ phase behaviour (Fig.4.25).

¹³ see Section 4.2.4.

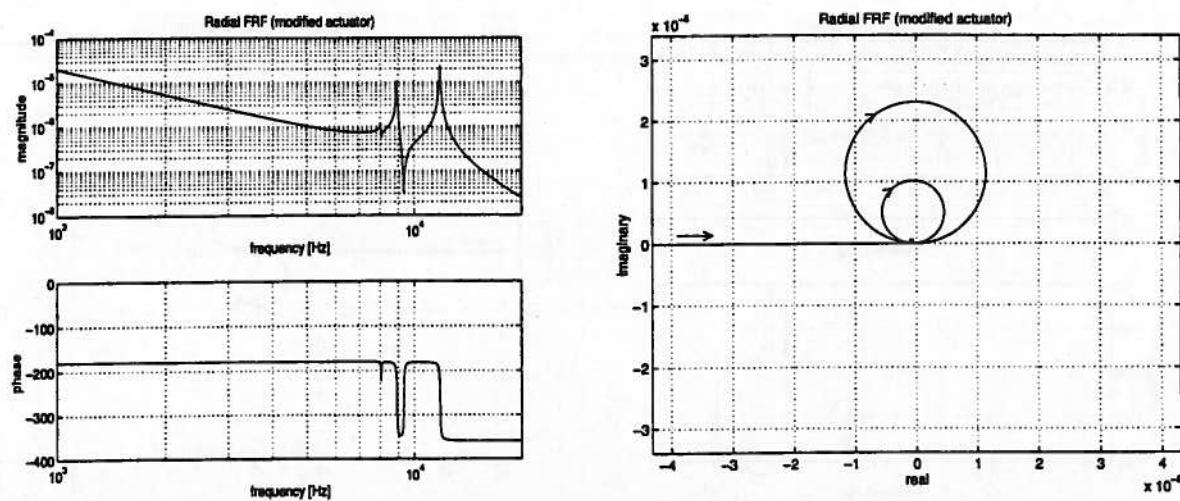


Figure 4.25 Radial FRF $x_{\text{lens}}/F_{\text{radial coil}}$ of modified actuator (left: Bode diagram, right: Nyquist diagram)

The mode-shape of the 8.8kHz mode is shown in Fig.4.26. As already expected from the modified phase behaviour, the radial movement of the coils and that of the lens are now in opposite directions, which leads to a negative value of α (see equation (4.3)) and thus to a pole/zero combination in the FRF (see Fig.4.5). Due to the phase lag that is present in the frequency range of interest, this newly found behaviour gives good stability performance, and no extra iterations were carried out.

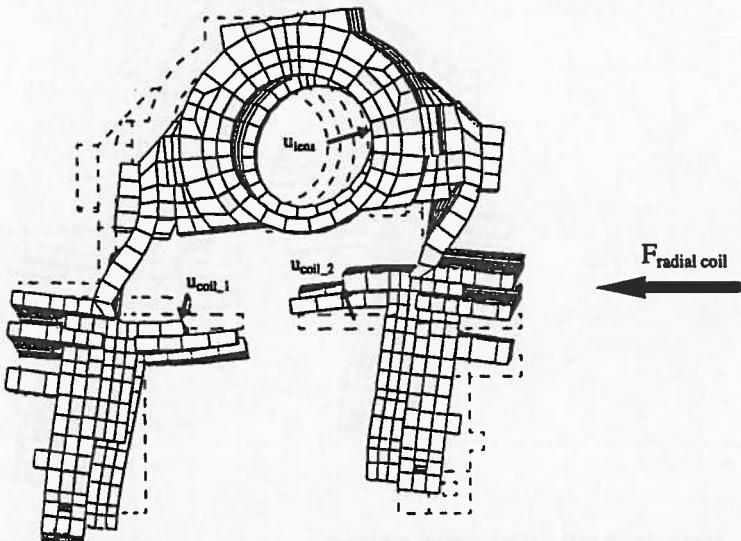


Figure 4.26 Mode shape at 8.8 kHz (modified actuator)

It is evident that as the original design resulted in a positive value of α , and the modified design in a negative value of α , it is also possible to find a modification which results in zero radial motion of the radial coils for that specific mode. Consequently, the mode would not be excited and therefore it would not be visible in the FRF.

The modelling stage will be followed by validation. This is an iterative process where the model is compared with experimental results and refined until the desired level of accuracy is reached. The validation stage is also used to verify that the model is able to predict future behaviors correctly.

5. Predictive Modelling

After the previous chapters about dynamics and its interaction with the servo control-loop, this chapter will discuss the use of modelling and simulation as a tool in the product-creation process of positioning devices.

The chapter starts with a short discussion of the steps involved in every modelling activity. Then, in the second part, a step-wise modelling approach is suggested, based on the idea of using the appropriate tools and the proper amount of detail in the successive stages of the product creation process. The focus is not directed towards the mathematics of the modelling tools, but towards their industrial application, such that the design process benefits from their application.

The third part of this chapter addresses three practical modelling issues, namely, sub-structuring, mixed dynamics/control simulations, and the influence of modal truncation on simulation results.

5.1. Steps in a Modelling Activity

The terms “Modelling” and “Simulation” are often used as synonyms to indicate the process in which a mathematical model of a real system (or design of a system) is derived, and certain properties are calculated. For example, one could try to calculate the static deflection of a wooden shelf due to the weight of a person that is using this “bridge” to cross a river (Fig.5.1).

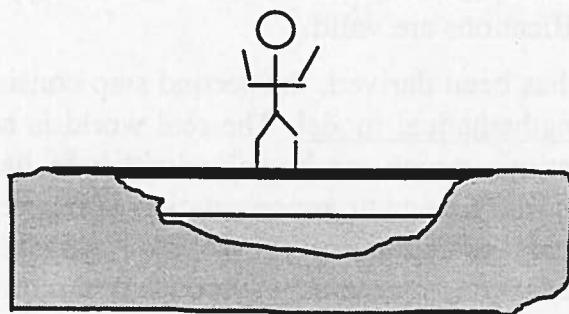


Figure 5.1 Static deflection of a wooden shelf over a river loaded by the weight of the user

Actually, modelling as a design tool should go further than the sole calculation of certain quantities; it should also include the interpretations of the calculation results. Including this final step, and starting with an initial design of a machine, one can distinguish at least four steps in any modelling activity (Fig.5.2).

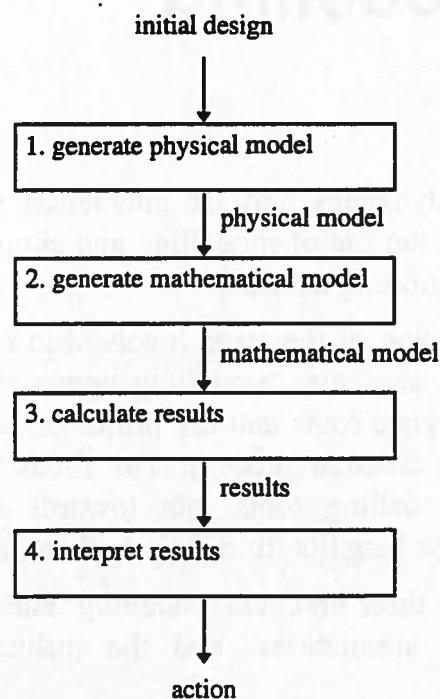


Figure 5.2 Steps in a modelling activity

The first step consists of a translation of the real structure or initial design drawing of a structure into a physical model. Such a physical model is a simplification of the reality, but contains all relations that are considered to be important to describe the investigated phenomenon. In the example, the wooden shelf could be represented by a beam with a uniform cross-section and uniform material properties, supported at its ends and loaded by a static force $F=m*g$. This step in the modelling process is the most difficult, because it requires experience and engineering judgement in order to determine, which simplifications are valid.

Once a physical model has been derived, the second step consists of translating this physical model into a mathematical model. The real world is now represented by a set of (differential) equations, which can be solved either by hand or by a computer program. In the example, the schematic representation of the beam is translated into a formula for the vertical deflection $x=Fl^3/(48EI)$. This second step is fairly straightforward, because it is based on existing approaches and rules.

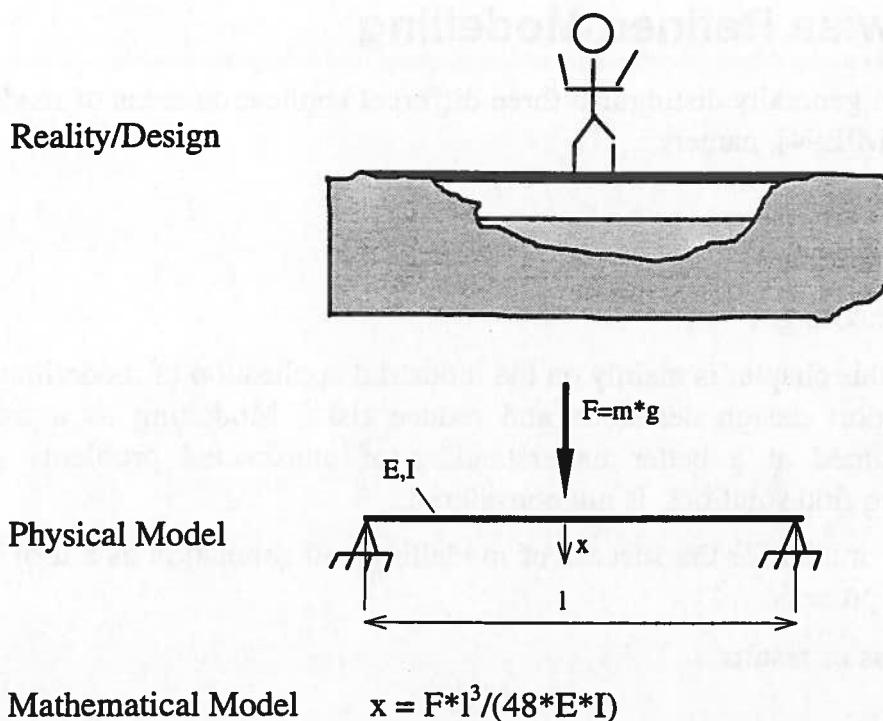


Figure 5.3 Illustration of the first two steps in the modelling process

The third step consists of the actual simulation run, the outcome of which is the value of some quantity (for example the deflection or stress in the wooden shelf) or its distribution as a function of time, place, or frequency.

The final step - interpretation of results - is really the most important in the whole process. Here, the calculated results and previously defined specifications are compared. On the basis of this comparison, design decisions are taken; for if the specified maximum deflection of the wooden shelf (distance between shelf and water) is exceeded, the user will get wet feet and consequently could decide to use a thicker shelf. It is important to realise that the design decisions taken in this step are the actual outcome of the modelling process.

Modelling and simulation can only have an impact on the design process when the last step is properly done. Often, a lot of time and energy is wasted because extended modelling and simulation is done with great enthusiasm only to find out at the end that nobody is capable of interpreting the results and to take design decisions on the basis of the obtained results. It is therefore recommended to start any modelling process by specifying the criteria that will be used in the interpretation and evaluation phase.

5.2. Step-wise Refined Modelling

Whereas one can generally distinguish three different application areas of modelling and simulation [MIL94], namely :

- Decision support
- Design optimisation
- Trouble shooting

the emphasis in this chapter is mainly on the industrial application of modelling tools in order to support design decisions and reduce risks. Modelling as a troubleshooting tool aimed at a better understanding of unexpected problems which hopefully helps to find solutions, is not considered.

Two aspects are crucial for the success of modelling and simulation as a tool in the product creation process:

- Usefulness of results
- Speed

The analysis program (and the analyst) must be capable of providing useful results, that is to say the answers to the proper questions. Simply stating that the first resonance frequency of a machine lies at 150Hz does satisfy the needs of the control engineer who wants to know whether the machine dynamics could endanger the stability of the servo system. It is important that the analysis program is capable of providing the data in the same sort of format that the people are used to. In the case of the servo-stability problem, the control engineer in most cases is used to evaluating Bode or Nyquist diagrams, therefore the results of the simulation must be presented in this format.

The second critical success factor is the speed (throughput time) with which the simulation results are obtained. The decision-making process can only be affected if the analysis results are available on time. If the analysis results are not available at a certain milestone in the design process, the decision must be made without this extra piece of information.

These considerations have led to an analysis and modelling approach that is based on the following two principles :

- Quick rejection of unsuitable proposals
- Balance between accuracy and speed

These principles can be achieved by an analysis process that is divided into distinct phases, which are closely related to, and go hand-in-hand with, the design process. Modelling and simulation are intended as a tool in the decision-making process, so it

is essential that they are done simultaneously, and that the type of simulation is adapted to the current phase in the development process (Fig.5.4).

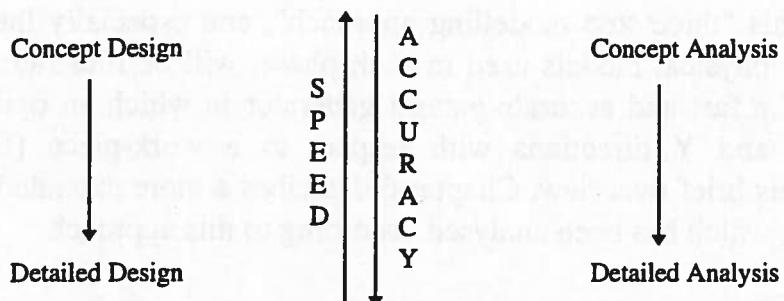


Figure 5.4 Phases in the design and modelling process

During the design process, which generally starts with various concept designs and finishes with one detailed design, the amount of information and detail is growing constantly. Concurrently, starting with very simple elementary models to support the selection of the proper concept, the simulation models become more refined during this top-down approach, just like the product under development. The modelling process must be preceded by a proper specification phase, in which the functional machine specifications are translated into servo- and dynamics-related specifications. These specifications will then be used to judge the results of the simulations.

In the past ten years, in which several projects have been completed, a three-step modelling approach has evolved, in which the following phases can be recognised :

- Concept Evaluation
- System Evaluation
- Component Evaluation

In concept evaluation, one checks whether a concept would work in 1D on the basis of a limited number of lumped masses connected by springs. In system evaluation, one checks whether it still works in 3D, again assuming rigid components connected by springs. Finally, in component evaluation one checks the deformation of the individual components, and how this affects the overall behaviour.

Opponents of computer simulation often doubt the predictive value of these simulations - especially if the models are very elementary - and therefore do not carry out these simulations. One must agree that successful simulations based on a 2-DOF lumped-mass model of a compact disc player or a chip-mounting machine are no guarantee that the final product will work according to specification. However, if these simulations, based on an elementary model of the product, show that the specifications are not met, then chances are extremely small that the final product will perform according to specifications. Therefore, computer simulations should be regarded as a means to guide the

design process by supporting the design choices and - which is probably the most important benefit - to detect unfit design concepts at a very early stage in the design process.

The philosophy of this "three-step modelling approach", and especially the typical level of detail of the physical models used in each phase, will be illustrated by the example (Fig.5.5) of a fast and accurate pattern generator in which an optical unit has to move in X and Y directions with respect to a work-piece [ELJ89/1]. Complementary to this brief overview, Chapter 6 describes a more extended case of a compact disc drive, which has been analysed according to this approach.

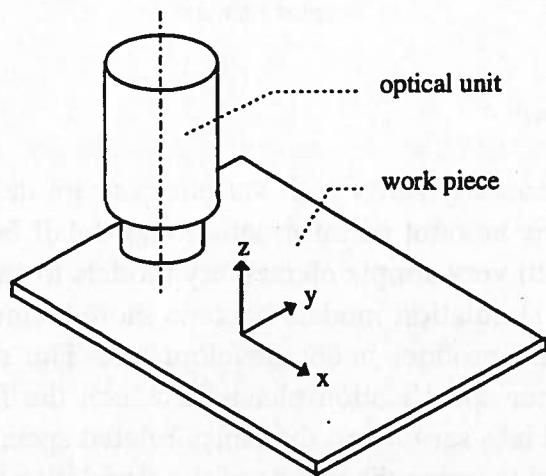


Figure 5.5 The basic elements of the pattern generator

The basic elements of this machine are the work-piece and the optical unit. The relative motion of these two elements in X and Y direction enables the generation of any pattern on the work-piece. Based on the required throughput of the machine, an acceleration level of 1m/s^2 is required, whereas the positioning accuracy is $1\mu\text{m}$ or better.

Specification

One of the most crucial, and at the same time most difficult, steps in the modelling process is the definition of proper criteria on the basis of which the simulation results can be judged. All too often, much time and money is wasted on simulations (the same is true for experiments) only to find out after completion that nobody is able to interpret the generated data and judge whether the results are satisfactory, or whether the design should be modified and, if so, how.

Before starting any simulations it is vital to derive proper criteria to judge the results. In most cases this implies that functional system-specifications in combination with assumed imperfections and disturbances need to be translated into dynamics and control specifications (Fig.5.6)

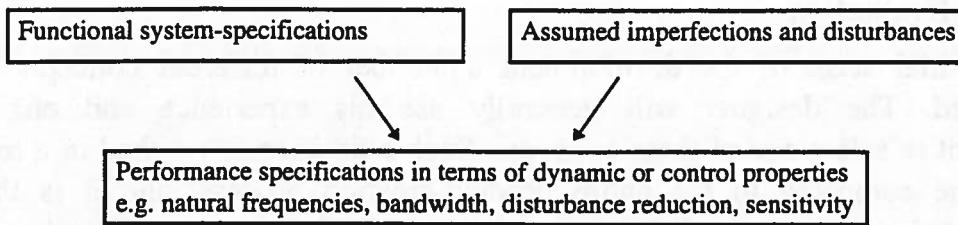


Figure 5.6 System performance specifications need to be translated into criteria on the basis of which simulation results can be judged.

In a first step one needs to make some initial estimation about the required bandwidth of the controlled system, because this is a prerequisite for evaluating the influence of the dynamics of the mechanical system on servo stability.

In the case of the pattern generator, one could aim at maintaining the desired accuracy of $1\mu\text{m}$ during the entire set-point profile, because this will avoid extra settling time. Not relying on settling time further implies that there is no time for an integrator action to compensate for static friction forces. Depending on the chosen configuration, friction forces up to 5-10N are realistic. To achieve the required accuracy during the prescribed motion one would typically have to use a controller with acceleration feedforward (see Section 2.1). However, any imperfection in the gain of this feedforward is similar to a disturbance force that needs to be compensated for by the feedback loop. With a set-point acceleration of 1m/s^2 and an assumed moving mass of 100kg, the nominal feedforward force amounts to 100N. For 5% feedforward mismatch, the feedback loop therefore has to compensate 5N.

On the basis of an assumed disturbance force (friction force plus feedforward mismatch) that equals 10N and the required accuracy of $1\mu\text{m}$, one can now make an initial estimate of the required servo stiffness k_p :

$$k_p = \frac{10 \text{ N}}{10^{-6} \text{ m}} = 10^7 \text{ N/m}$$

Neglecting the effect of the derivative action of the controller, one can obtain a first engineering estimate of the bandwidth f_b (equals the 0-dB crossing of the open-loop response) on the basis of the servo stiffness k_p and the moving mass m :

$$f_b \approx \frac{1}{2\pi} \sqrt{\frac{k_p}{m}}$$

With the previously calculated servo stiffness of 10^7N/m and a total moving mass of about 100kg, the required bandwidth amounts to 50Hz. Having derived this estimate of the required bandwidth on the basis of the necessary disturbance rejection, one has to consider whether this bandwidth can be achieved without introducing stability problems and what the consequences are for the mechanical design. On the basis of this consideration, the dynamic properties of the various designs can now be evaluated.

Concept Evaluation

In the initial stage of the development a number of different concepts will be considered. The designer will generally use his experience and engineering judgement to select one of these concepts. Such a decision is reached in a relatively short time compared to the entire product-creation process, and it is therefore essential to have results from dynamic simulation available within a few days. In this stage the designer only has a rough idea about the outlines of the machine, and the feasibility of this idea can be judged on the basis of very elementary calculations.

One of the potential concepts for this machine consists of a stationary work-piece with an optical unit that moves in both the X and Y direction (Fig.5.7). In the X direction two driving forces are applied to the slides, whereas the position is measured by two linear encoders mounted between the slide and the granite frame.

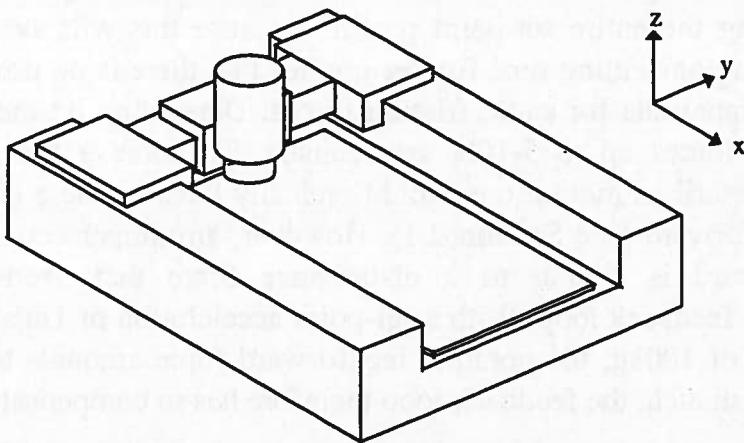


Figure 5.7 One of the possible concepts of the pattern generator

In this stage of the design a simple model of the dynamic effects in the X direction could consist of the base, the slides, the guiding rail, the optical housing, and intermediate flexibility (Fig.5.8). This simple symmetrical model is only valid as long as the optical unit is in the middle of the Y travel. For other positions the model should be expanded to include extra rotational DOF.

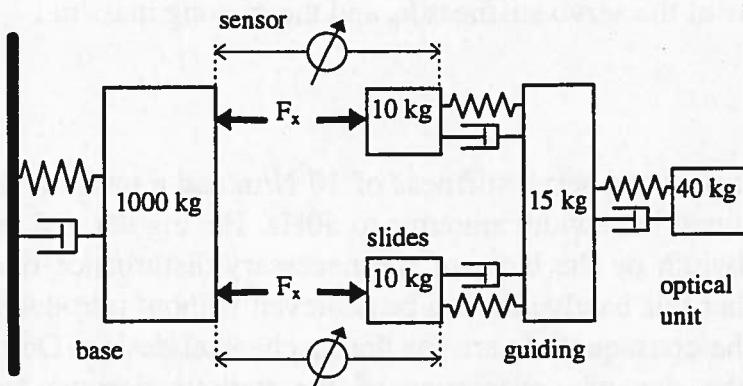


Figure 5.8 Simple 1D model for the analysis of the dynamic behaviour in X direction

By using this fast and simple method of analysis, potential risks associated with the different concepts can be evaluated. These results, together with all other design criteria, are then used to select a feasible concept.

System Evaluation

Once the concept of the machine has been chosen, the first rough three-dimensional sketches become available and one can add extra spatial information to the simulations, such as :

- mass and mass moment of inertia of the different components
- location of the centre of gravity
- location of connecting stiffness
- location of driving force(s)
- location of sensor(s)

Typically, such a model contains 5-10 rigid bodies connected by suitable connectors that incorporate flexibility, whereas damping is in most cases added in the form of modal damping (1% relative damping is in most cases a good first estimate if no special attempts have been made to achieve extra damping).

Figure 5.9 shows such a 3D model of a different concept for the pattern generator. In this concept the X and Y motions are split between the mounting head, which can only move in the Y direction, and the work-piece, which is mounted on a linear motor that can move in the X direction.

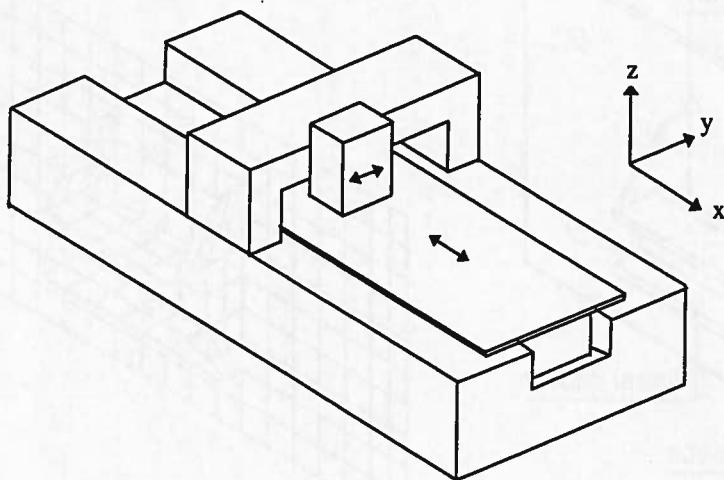


Figure 5.9 Rigid Body Model of a concept based on a movement of the work-piece in X direction, and a movement of the optical unit in Y direction.

Component Evaluation

On the basis of previous analyses, experimental evaluation of previous designs, or engineering judgement, it is generally possible to identify critical components in the design. These components will then need to be analysed in more detail using FEM.

Sometimes it is possible to judge the influence of the internal dynamics of such a component on the performance of the total system, based on a separate analysis of the component. This approach requires serious consideration of the boundary conditions and the translation of system-performance requirements into requirements on the basis of which a component can be evaluated.

Such a separate analysis of a component is not always feasible, in which case the detailed FEM description of the component can be used to replace the former rigid-body description that has been used in the "System Evaluation". Such a step normally requires the use of so-called "sub-structuring" techniques, the discussion of which is part of Section 5.3.

In the pattern generator it is very important that the connection between the linear motor module and the work-piece is sufficiently stiff. The reason lies in the fact that due to accuracy specifications the position is measured at the work-piece (using laser interferometers) and not at the motor. Consequently, one has to ensure that the internal stiffness of the actuator is high enough to avoid the stability problems discussed in Section 4.3.1. In one of the concepts a set of thin plates were used to obtain a proper connection. This idea is shown in Fig.5.10, together with the finite element mesh.

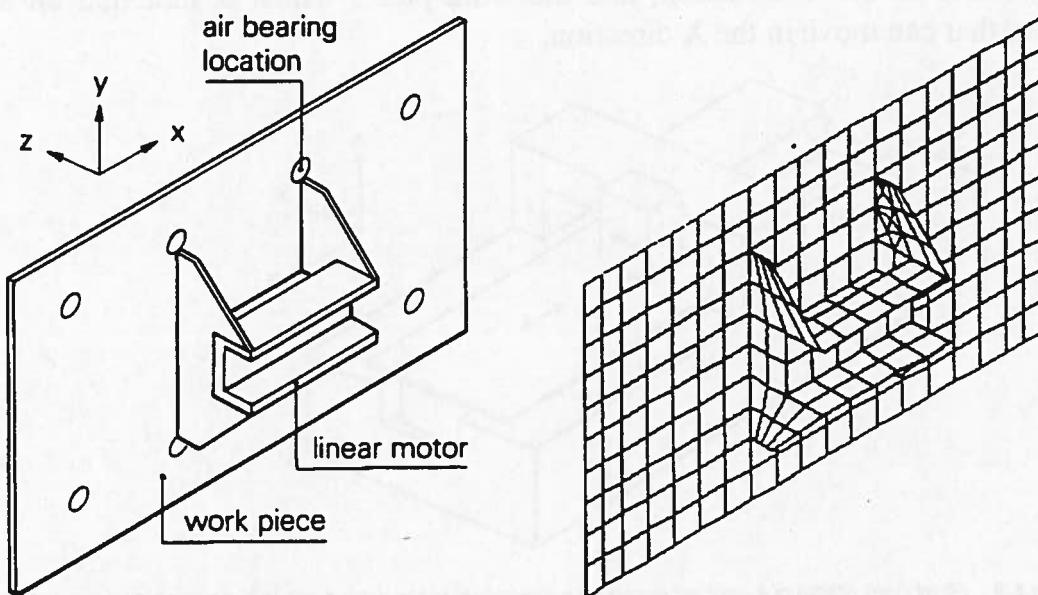


Figure 5.10 Construction using thin plates to connect the linear motor module to the work-piece (sketch and finite element mesh)

Final Remarks

For the industrial application of “Predictive Modelling” it is essential that the amount of detail in a simulation model corresponds to the current phase in the design process. A design team profits from the application of simulation tools only if a proper balance is found between detail and accuracy on one hand, and the total throughput time of the analysis on the other hand. In the context of using “Predictive Modelling” as a tool to support design decisions, a prediction of resonance frequencies of a mechanical system with an accuracy of 25% is considered good enough. This is quite different from the approach used in the field of design optimisation, in which one tries to achieve a better agreement by adapting the model on the basis of measured dynamic properties.

5.3. Practical Modelling Issues

In the previous discussion of the step-wise modelling philosophy, no attention was paid to a number of practical modelling issues. The essence of these practical considerations is :

1. How to deal with large detailed models of the entire system
2. How to perform simulations that include the mechanical system plus the controller

Ad 1

In the “Component Evaluation” stage of the development and modelling process, detailed FEM models of critical components need to be created and analysed. Sometimes, components can be evaluated individually against component specifications, which are often defined in terms of lowest internal natural frequencies. In other cases, such a separate analysis of a component is not sufficient to judge the impact of its dynamics on the overall system, and one is forced to combine these detailed component-models into a detailed model of the entire system.

Due to the complexity of the structures it is normally not very practical to build one, single, huge, FEM of the entire device. Such an approach has serious limitations :

- Building one huge model of a machine tends to be very error-prone.
- It is not feasible to work on one huge model with a group of people.
- The resulting mass and stiffness matrices can easily have many thousand degrees of freedom, which puts high demands on the required computing capacity. Even in the computer age, this effect should not be under-estimated (several hours or even days analysis time on a standard CAD workstation are no exception).

- Furthermore, in the philosophy of constant interaction between the creative design process and the analytical evaluation, many iteration cycles are required. In this stage of the development, these iterations should normally focus on modifications of individual components of the system, and it is very inefficient and time-consuming if any change of one single component requires the re-analysis of the entire system without re-using previous results.

A technique which overcomes these disadvantages is the so-called "Sub-structuring Technique". In this approach, illustrated in Fig.5.11, the system is divided into sub-structures or components, which are analysed separately. Then, the (reduced) models of the components are assembled to form the overall system. By doing so, the size of the final system model is significantly reduced.

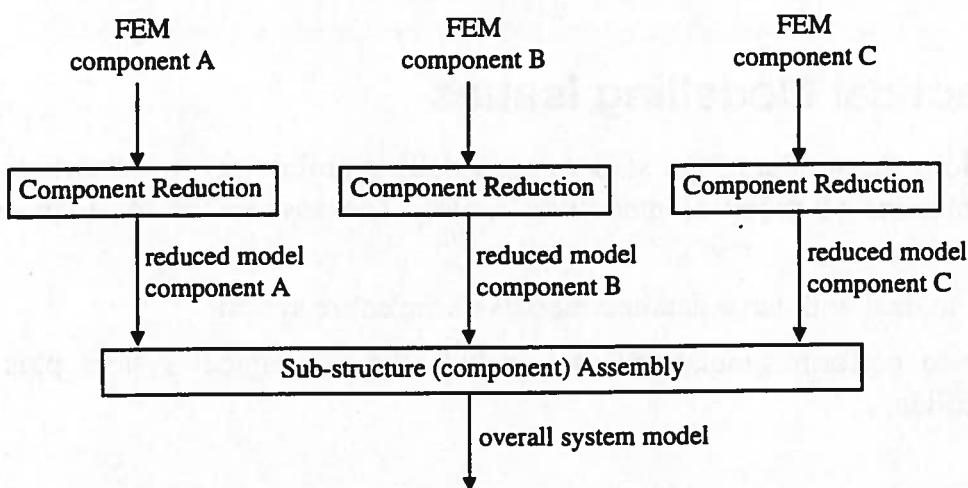


Figure 5.11 Steps in the creation of an overall system model based on detailed FE-models of the components

The process involves the following steps :

- The first step, the subdivision of the entire system into components, is normally very straightforward, and very much driven by the physical modules and interfaces of the machine.
- In the second step a detailed FE-model of each component is generated, resulting in a component mass and stiffness matrix.
- In step three, a reduced model of the component is generated by applying a "component reduction" technique to the original model. The intention of this step is to reduce the size of the matrices that describe the behaviour of the component, yet retain its main dynamic characteristics.
- Finally, the reduced models are assembled into one overall system.

Ad 2

Depending on the type of analysis, one can restrict the analysis entirely to the dynamics of the mechanical device, or one needs to include the servo control-system and do a mixed dynamics/control analysis. The way in which these combined dynamics-control simulations are done depends very much on the functionality of the software that is being used in a certain stage of the modelling process.

As long as the mechanical model is restricted to a number of lumped one-dimensional or two-dimensional rigid bodies connected by springs and dampers, one can use one single modelling tool, such as Simulink/Matlab or SPICE-based network analysis programs, to model both the mechanics and the control system. However, as soon as one needs to make detailed models of the mechanical system, one has to apply FEM programs, which normally do not include standard features for modelling control systems.

Some FEM programs support user-defined subroutines, or a sort of macro-language, which potentially enables the addition of control systems¹. However, this approach requires software coding and is considered very laborious and error-prone. These considerations have led to adapting another approach (Fig.5.12).

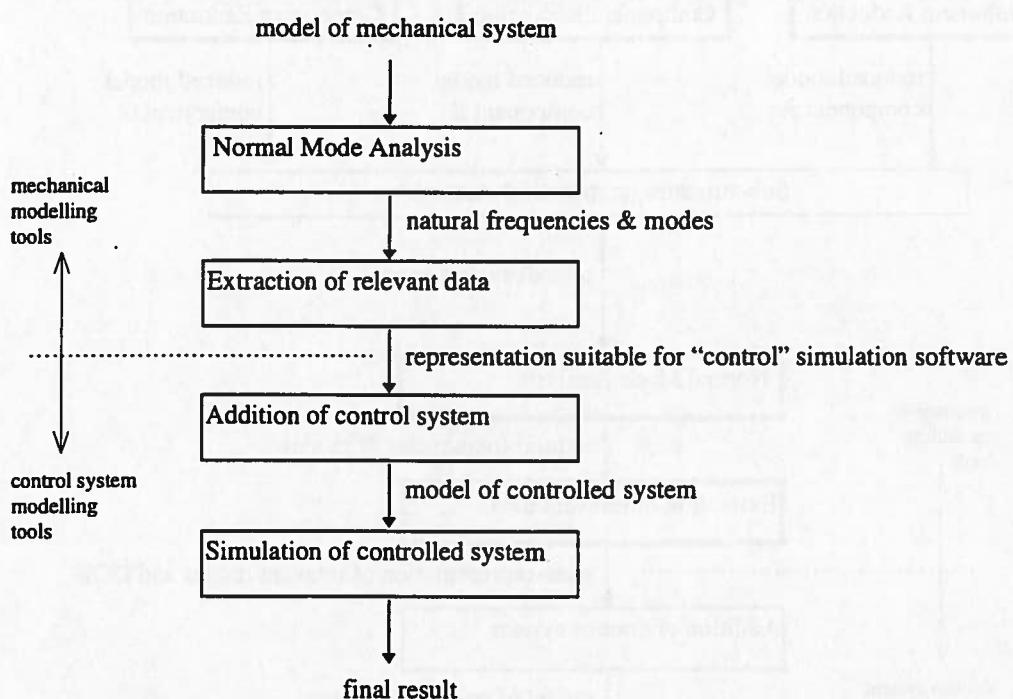


Figure 5.12 Steps in the creation of an overall system model that includes the mechanics plus the control system

¹ Gielens has shown in his master thesis [GIE92/1] that a reasonably user-friendly control module can be integrated into the dynamics software package I-DEAS/System Dynamics on the basis of a combination of internal program files and external C code.

It is based on exporting the dynamic characteristics of the mechanical system in a suitable form from the mechanical analysis software to some other modelling program in which the control system can be added, and in which the overall behaviour can be simulated. Typically, the reduced model of the dynamic properties of the mechanical system is based on a truncated set of modes and modal parameters, but other techniques are also used [WOR94].

By combining Figs.5.11 and 5.12, a complete overview of the various steps required for the detailed modelling of a controlled positioning device can be obtained (Fig.5.13). Each individual step must be done with great care in order to ensure the integrity of the final result. The assembly of mechanical components to create an overall model (see Appendix E), and the steps after the transfer to a simulation environment in which the control system can be added, are considered basic operations. The three steps indicated by bold lines, and some related topics, will be discussed in the remainder of this chapter.

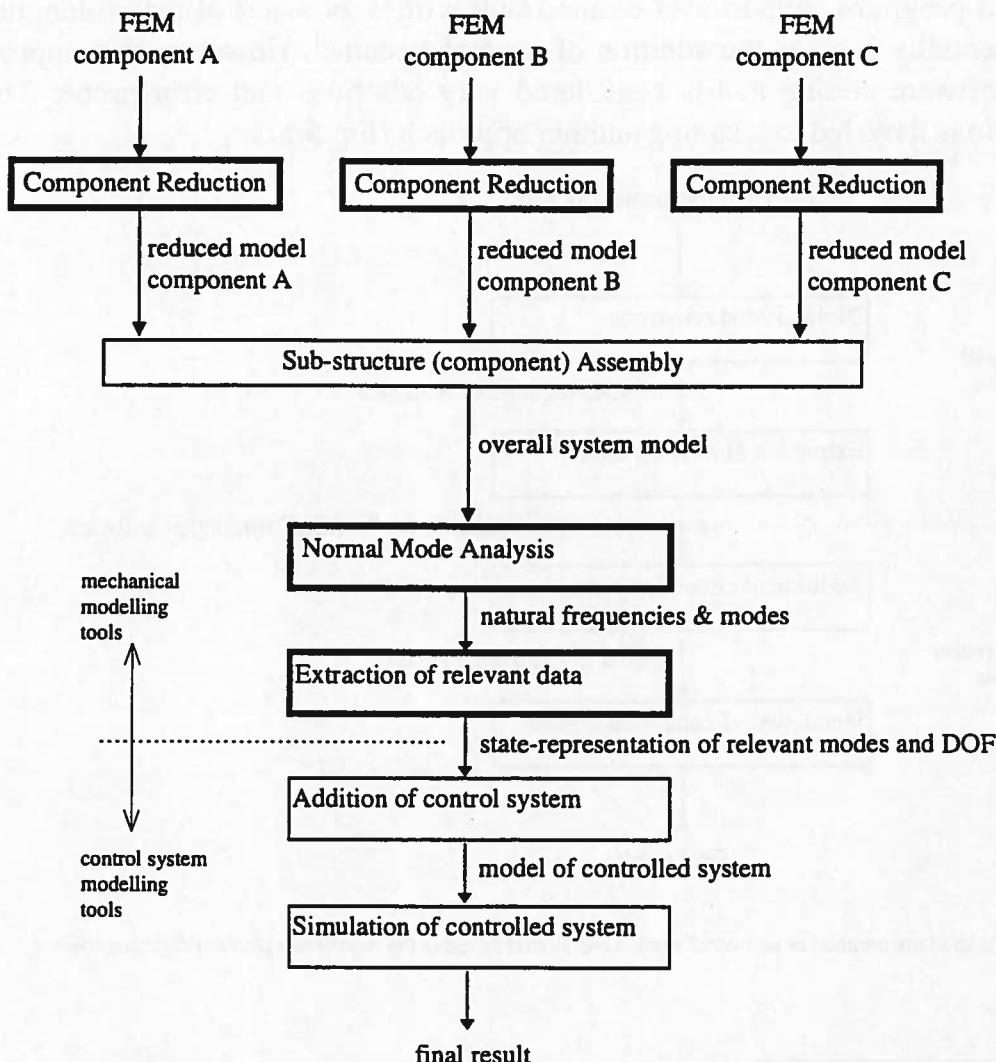


Figure 5.13 Complete overview of steps required for the simulation of a servo-controlled device that includes a detailed model of the mechanical system (various FEM components) plus the control system

5.3.1. Mixed Dynamics/Control Simulations

In many cases, time-domain simulations of a manipulator, including its servo controller, cannot be done with the modelling tools used for the analysis of the mechanical system. As a result, the dynamic properties of the mechanical system need to be transferred to some other modelling tool which is capable of adding the control system. Generally, this transfer is based on the modes and modal parameters of the mechanical system. Assuming that no (animated) visualisation of the entire system is required, only the mode-shape data of the following physical DOF are relevant :

- excitation DOF : DOF where forces (driving forces, disturbing forces) are applied or motion (floor vibration) is prescribed.
- response DOF : DOF which are used in the feedback loop (servo position) or which are controlled (distance between tool and payload) or which need to be monitored (e.g. frame motion)

Assuming that the number of relevant DOF is equal to r , and the number of transferred modes is equal to m , the dynamic properties can be characterised by m sets of modal parameters m_i , k_i , and mode-shape vectors ϕ_i of length r . Damping can be added in the form of modal damping $d_i = 2\beta_i m_i \omega_i$, in which β_i represents the relative modal damping (typical values are in the order of 1%) :

$$\begin{vmatrix} x_1 \\ x_2 \\ \vdots \\ x_r \end{vmatrix} = q_1 \begin{vmatrix} \phi_{11} \\ \phi_{12} \\ \vdots \\ \phi_{1r} \end{vmatrix} + q_2 \begin{vmatrix} \phi_{21} \\ \phi_{22} \\ \vdots \\ \phi_{2r} \end{vmatrix} + \dots + q_m \begin{vmatrix} \phi_{m1} \\ \phi_{m2} \\ \vdots \\ \phi_{mr} \end{vmatrix} \quad (5.1)$$

$$\begin{bmatrix} m_1 & & \ddot{q}_1 & \begin{bmatrix} k_1 & & q_1 & \begin{bmatrix} \phi_{11} & \phi_{12} & \dots & \phi_{1r} \end{bmatrix} f_1 \\ m_2 & & \ddot{q}_2 & \begin{bmatrix} k_2 & & q_2 & \begin{bmatrix} \phi_{21} & \phi_{22} & & \phi_{2r} \end{bmatrix} f_2 \\ \vdots & & \vdots & \vdots & \vdots & \vdots \\ m_m & & \ddot{q}_m & \begin{bmatrix} k_m & & q_m & \begin{bmatrix} \phi_{m1} & \phi_{m2} & \dots & \phi_{mr} \end{bmatrix} f_r \end{bmatrix} \end{bmatrix} \end{bmatrix} \quad (5.2)$$

On the basis of the modal lever approach discussed in Chapter 3, procedures have been derived [RAN88/4,RAN88/5] to transfer this set of data to SPICE-based electronic network analysis software, which can be very profitable in simulations of controlled manipulators. However, in popular control-oriented programs such as Matlab/Simulink, the description of a system is based on a state-variable representation, which has the standard form :

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx + Du \end{aligned} \quad (5.3)$$

and therefore equations (5.1) and (5.2) must be adapted to that representation.

Considering the modal DOF q_i and modal velocity \dot{q}_i as the states of the system, the modal state representation (5.4) can be derived, in which forces can be applied at all included physical DOF, and the response of all physical DOF can be monitored. Damping has been added in the form of relative modal damping β_i .

$$\begin{bmatrix} \dot{q}_1 \\ \dot{q}_2 \\ \vdots \\ \dot{q}_m \end{bmatrix} = \begin{bmatrix} 0 & & & 1 & & \\ & 0 & & & 1 & \\ & & \ddots & & & \ddots \\ & & & 0 & & \\ & -\omega_1^2 & & -2\beta_1\omega_1 & & 1 \\ & & -\omega_2^2 & & -2\beta_2\omega_2 & \\ & & & & & \ddots \\ & & & -\omega_m^2 & & -2\beta_m\omega_m \end{bmatrix} \begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_m \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ m_1^{-1}\phi_{11} & m_1^{-1}\phi_{12} & \cdots & m_1^{-1}\phi_{1r} \\ m_2^{-1}\phi_{21} & m_2^{-1}\phi_{22} & \cdots & m_2^{-1}\phi_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ m_m^{-1}\phi_{m1} & m_m^{-1}\phi_{m2} & \cdots & m_m^{-1}\phi_{mr} \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_r \end{bmatrix} \quad (5.4)$$

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_r \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{21} & \cdots & \phi_{m1} \\ \phi_{12} & \phi_{22} & \cdots & \phi_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{1r} & \phi_{2r} & \cdots & \phi_{mr} \end{bmatrix} \begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_m \end{bmatrix}$$

If only one physical response or output DOF x_o , and one excitation DOF x_e exists, equation (5.4) is reduced to :

$$\begin{bmatrix} \dot{q}_1 \\ \dot{q}_2 \\ \vdots \\ \dot{q}_m \end{bmatrix} = \begin{bmatrix} 0 & & & 1 & & \\ & 0 & & & 1 & \\ & & \ddots & & & \ddots \\ & & & 0 & & \\ & -\omega_1^2 & & -2\beta_1\omega_1 & & 1 \\ & & -\omega_2^2 & & -2\beta_2\omega_2 & \\ & & & & & \ddots \\ & & & -\omega_m^2 & & -2\beta_m\omega_m \end{bmatrix} \begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_m \end{bmatrix} + f_e \begin{bmatrix} m_1^{-1}\phi_{1e} \\ m_2^{-1}\phi_{2e} \\ \vdots \\ m_m^{-1}\phi_{me} \end{bmatrix} \quad (5.5)$$

$$x_o = \begin{bmatrix} \phi_{1o} & \phi_{2o} & \cdots & \phi_{mo} \end{bmatrix} \begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_m \end{bmatrix}$$

which contains as few as $5*m$ parameters ($\omega_i, m_i, \beta_i, \phi_{ie}, \phi_{io}; i=1,m$) to characterise the entire behaviour of this configuration.

On the basis of the state-variable representation, an automated procedure has been derived to extract and export the relevant dynamic properties from an I-DEAS/System Dynamics model file to Matlab/Simulink [HEI94].

5.3.2. Effect of Modal Truncation and Accuracy

In most cases, the description of the dynamic behaviour of the mechanical system is based on its modes and modal parameters. When using this approach, one has to be aware of unexpected side effects caused by :

- modal truncation
- limited accuracy of the mode-shape vectors

In almost all practical applications, the user is only interested in a limited frequency range. In the stability analysis of a servo control-loop, the maximum frequency of interest is generally limited to 10-20 times the desired servo bandwidth, whereas the effect of floor vibrations on the performance of a system is dominated by the modes below 100-200Hz. Combined with the desire to reduce analysis time, one will normally apply “modal truncation”, and include only those modes in the analysis that fall within the frequency range of interest. However, it is important to be aware of the fact that modal truncation violates the static solution capability².

According to equation (3.16) the transfer function (x_i/F_k) can be written as :

$$\left(\frac{x_i}{F_k} \right) = \sum_{i=1}^n \left(\frac{x_i}{F_k} \right)_i = \sum_{i=1}^n \frac{\phi_{ik} \phi_{il}}{m_i s^2 + k_i} \quad (5.6)$$

It is evident that this relation must also be true for the static situation, and therefore one can write for the static displacement of x_i as the result of a static force f_k :

$$\left(\frac{x_i}{f_k} \right)_{\text{static}} = \sum_{i=1}^n \left(\frac{x_i}{f_k} \right)_{i,\text{static}} = \sum_{i=1}^n \frac{\phi_{ik} \phi_{il}}{k_i} \quad (5.7)$$

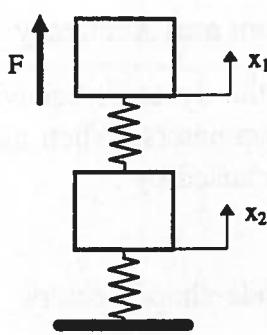
As long as all n modes of the system with n physical DOF are included in the modal superposition, the result will be identical to that obtained from a direct solution of the static problem. However, if only m modes are included, and $(n-m)$ high-frequency modes are excluded, the result of the static analysis will be different because the contributions of the truncated modes to the static solution are not taken into account (see equation 5.8)).

$$\left(\frac{x_i}{f_k} \right)_{\text{static}} = \sum_{i=1}^m \frac{\phi_{ik} \phi_{il}}{k_i} + \sum_{i=m+1}^n \frac{\phi_{ik} \phi_{il}}{k_i} \quad (5.8)$$

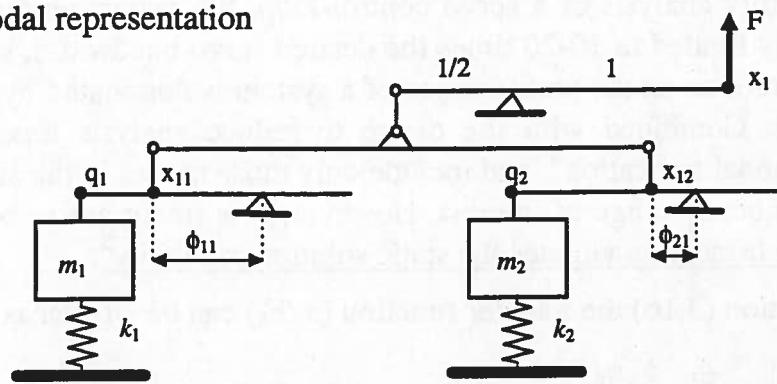
This effect is easily demonstrated graphically when considering a system with only two DOF (Fig.5.14a), which can be fully described by its two modes, as indicated in Fig.5.14b. Neglecting the influence of mode 2, which is the same as adding an extra fixation as indicated in Fig.5.14c, will yield different static results when applying a static force F at physical DOF x_1 .

² A reduced model that possesses static solution capability yields the same (quasi-)static solution as the original, unreduced model.

a. original system



b. full modal representation



c. truncated modal representation

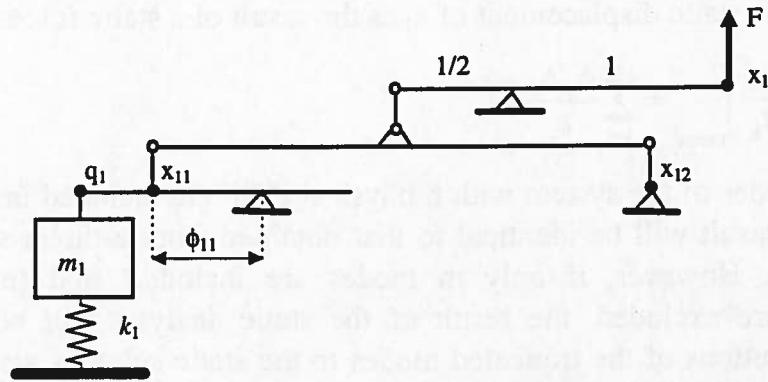


Figure 5.14 a. Original system b. Full modal description c. Truncated modal description

The effect of modal truncation is even more bizarre when looking at the relative displacement of $(x_1 - x_2)$ as result of a force F that is located at the physical DOF x_2 . Engineering judgement says that the resulting internal deformation $(x_1 - x_2)$ must be equal to zero, and a complete modal superposition yields that result. However, when having a closer look at the individual modal contribution to the static solution, one will always find two terms in this special case; these cancel each other (independent of the actual stiffness and mass distribution). Consequently, when excluding one of the two modes from the modal superposition, a non-zero static deformation $(x_1 - x_2)$ is found.

The fact that modal truncation neglects all modes above a certain frequency range, and that therefore the FRF will show no resonance peaks above the truncation frequency, is an accepted fact; however, the fact that the static solution capability is no longer guaranteed can lead to unexpected phenomena. As an illustration, two of these unexpected phenomena, which have been observed in practice, will be explained here :

- set-point response of a servo system with stiffness in motion direction
- sensitivity of a system to low frequency external vibrations

Set-point Response of a Servo System with Stiffness in Motion Direction

Fig.5.15 shows a schematic representation of a servo-controlled positioning device which has the following essential properties :

- the guiding system is such (for example a leaf-spring guiding of a compact disc actuator head) that in the motion direction a certain counteracting stiffness k is present.
- the performance-related displacement is the position of the payload relative to some reference ($x_{accuracy}$), whereas the feedback of the control system is based on measuring the position of the slide relative to the reference (x_{servo}).

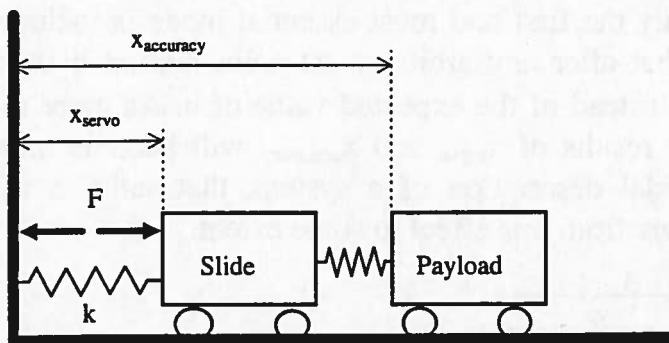


Figure 5.15 Schematic representation of servo system

In the time-response simulation of such a system, if it is based on a truncated modal description, an unexpected side effect can be observed when studying the response to a position set-point. Fig.5.16 shows the time response of the controlled system to a very gentle set-point profile of 10^{-2}m in 1 sec. When monitoring the measured position x_{servo} , one sees that, due to the integrator part in the PID controller, the desired set-point value is reached exactly, and the servo error equals zero. However, the position of the payload ($x_{accuracy}$) does not reach the desired set-point position, and a static-position error remains. In the absence of any non-linear effect, such as friction, backlash, and integrator limits, this effect is very puzzling but can be explained by the effect of modal truncation.

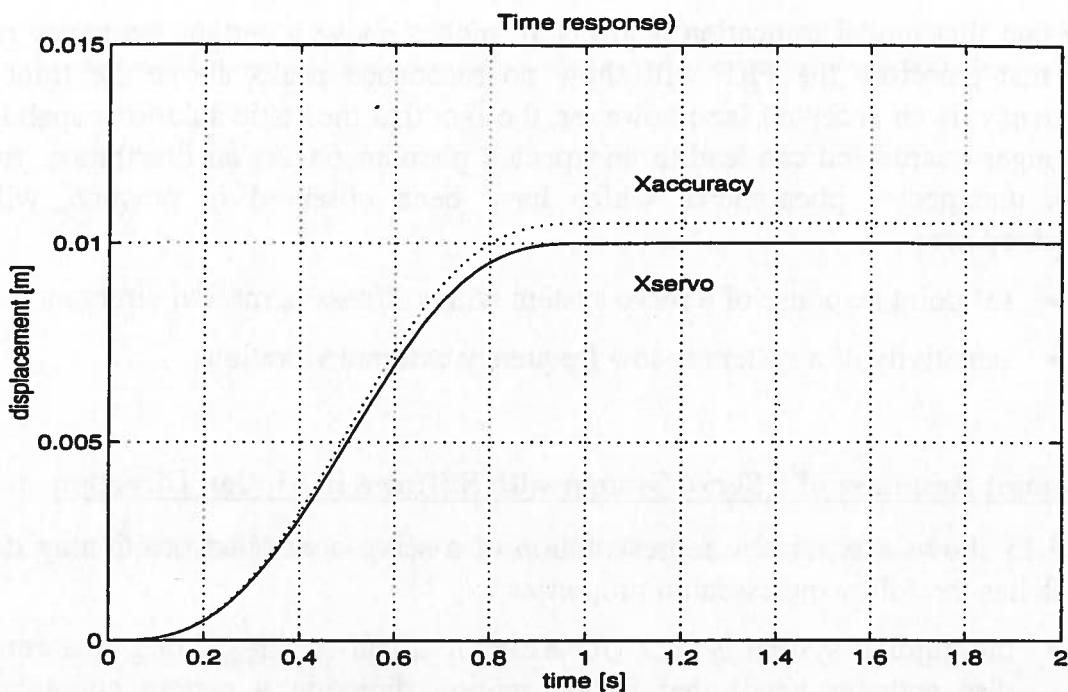


Figure 5.16 Time response of x_{servo} and x_{accuracy}

As a result of the stiffness k , although it is generally small, the first mode will not be a rigid-body mode, and consequently a small difference between the modal displacement of the slide ($x_{\text{servo}}=u$) and the payload ($x_{\text{accuracy}}=\gamma^* u$) will exist (see equation (5.17)). Thus, if only the first and most essential mode is included in the analysis, one will observe that after any arbitrary set-point motion h the distance x_{accuracy} will be equal to $\gamma^* h$ instead of the expected value of h . As more modes are included in the analysis, the results of x_{servo} and x_{accuracy} will become more equal. However, any truncated modal description of a system, that satisfies the above mentioned assumptions, suffers from this effect to some extent.

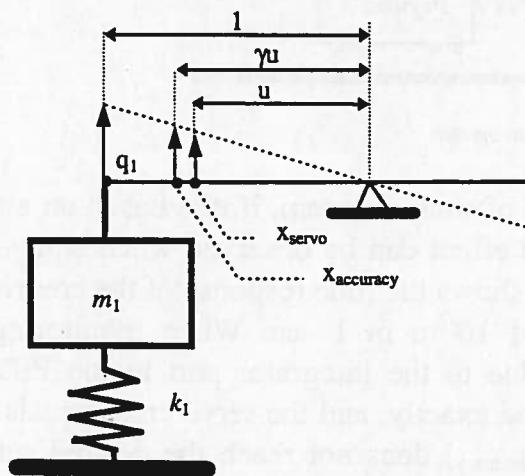


Figure 5.17 First mode of a system with stiffness in motion direction

Sensitivity of a system to low-frequency external vibrations

Figure 5.18 shows a 2-DOF model of a precision instrument such as an electron microscope or wafer-stepper. It consists of a heavy machine frame (mass m_2) relative to which a light manipulator (mass $m_1 \ll m_2$) can be positioned. The sensitivity of the device to external vibrations h , is characterised by the transfer function (ε/h), in which ε represents the internal error. To isolate the system from these vibrations it is supported by a very soft suspension k_{20} (typical suspension frequency³ $f_1 = 3\text{Hz}$), whereas the internal resonance³ of the machine f_2 is designed as high as possible ($>100\text{Hz}$).

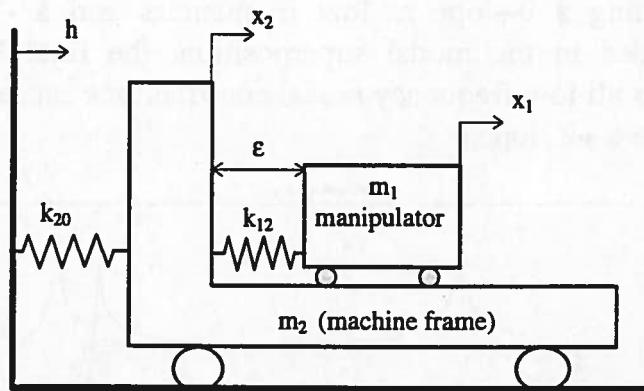


Figure 5.18 Lumped Mass Model of a Machine

In a lightly damped system, the transfer curve (ε/h) is typically characterised by a +2 asymptote until the suspension frequency, a -2 asymptote above the internal resonance f_2 , and a horizontal part in between, with level $(f_1/f_2)^2$. Figure 5.19 shows such a curve, assuming a suspension frequency of 3Hz (5% relative damping) and an internal mode at 100Hz with only little relative damping (1%).

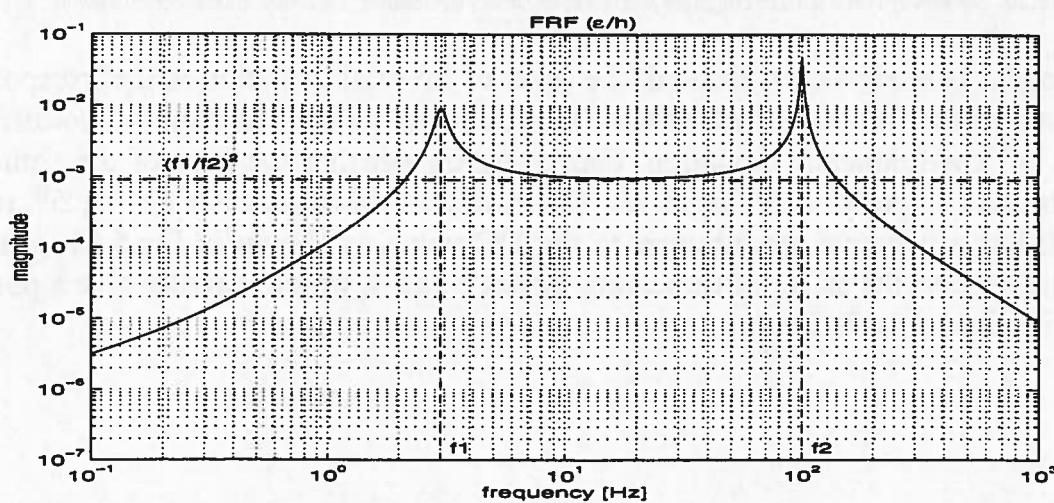


Figure 5.19 Typical Sensitivity Curve (ε/h)

³ $f_1 \approx \frac{1}{2\pi} \sqrt{\frac{m_1+m_2}{k_{20}}}$ $f_2 \approx \frac{1}{2\pi} \sqrt{\frac{m_1}{k_{12}}}$

An engineering explanation of the general shape of this curve is that at very low frequencies the sensitivity to floor vibrations should approximate zero, because a static motion (= zero Hz) of the floor, and consequently of the entire machine, does not lead to any internal deformation due to the absence of acceleration forces. At frequencies far above the internal resonance the machine frame is completely decoupled from the motion of the floor.

However, as soon as such a sensitivity curve is based on a truncated modal superposition, it can display significant deviations at low frequencies. This effect can best be illustrated by considering an imaginary system with five DOF. If all five modes, each contributing a 0-slope at low frequencies and a -2 slope at high frequencies are included in the modal superposition, the final FRF will be as expected (Fig.5.20). As all low-frequency modal contributions cancel each other, the final result will become a +2 slope.

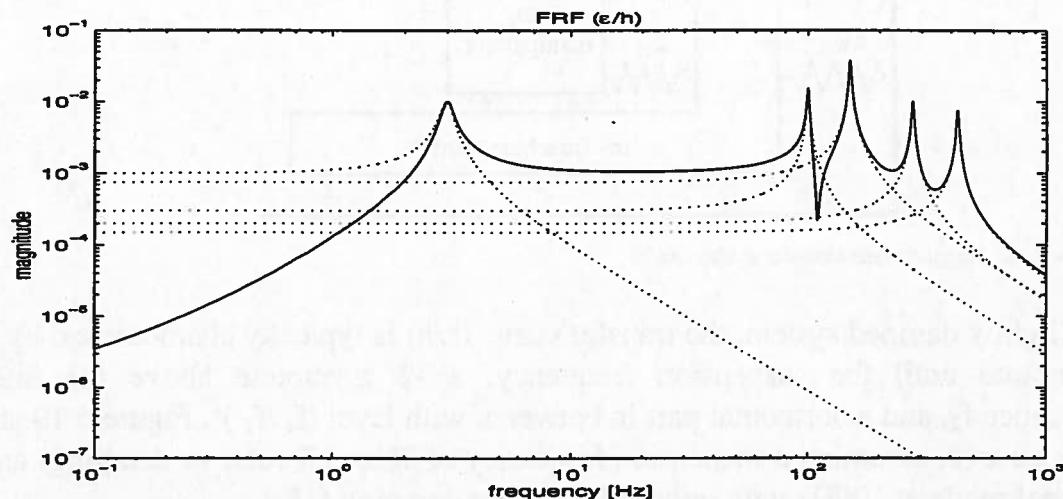


Figure 5.20 Sensitivity curve (ϵ/h) of complete model (solid : final FRF, dotted : individual modal contributions)

Evidently, if modes are truncated, the level of the static solution will correspond to the sum of the truncated flexibilities. Truncation of modes can lead to two different types of low-frequency deviation, depending on whether the sum of the truncated flexibilities is positive or negative. The result of the truncation of the 5th mode, which has a negative contribution to the DC value, is shown in Fig.5.21, whereas Fig.5.22 shows the effect of truncating modes 5 and 4, which together give a positive contribution to the DC value.

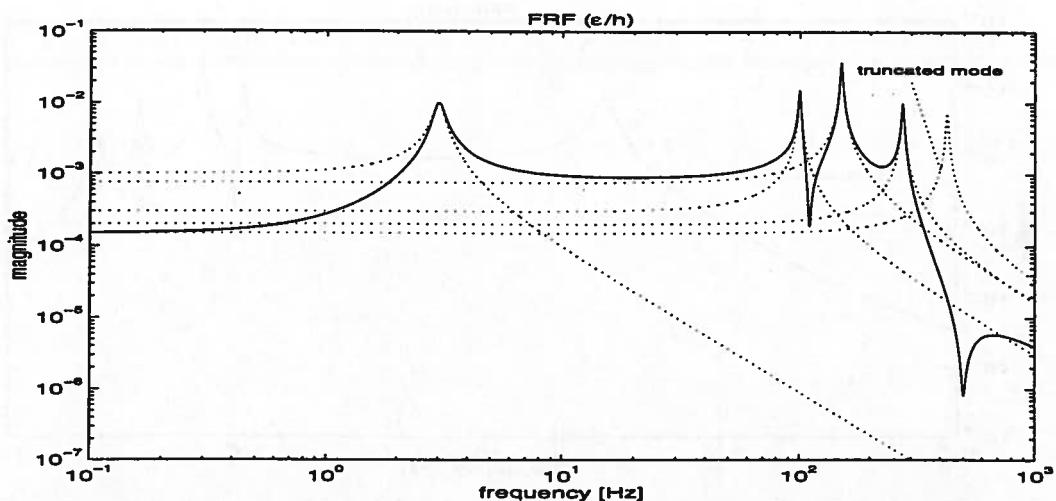


Figure 5.21 Sensitivity curve (e/h) after truncation of mode 5 (solid : final FRF, dotted :modal contributions)

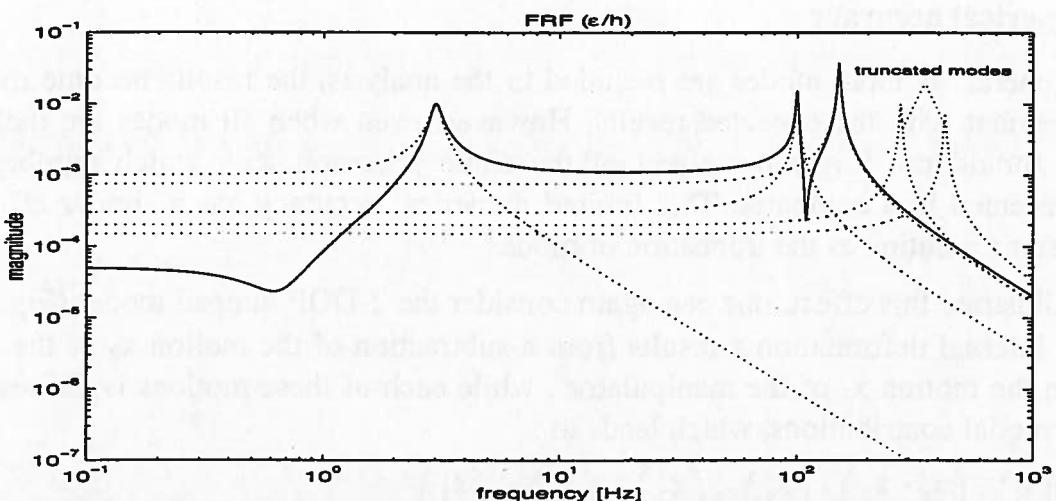


Figure 5.22 Sensitivity curve (e/h) after truncation of mode 4&5 (solid : final FRF, dotted : modal contributions)

How these two characteristic shapes emerge can best be explained by considering that excluding a mode from the modal superposition is identical to subtracting this mode from the complete solution. In Fig.5.23 the FRF corresponding to the complete solution plus the contribution of one individual mode is shown. As one can see, an intersection of these two curves is found below the suspension frequency. At this intersection frequency the amplitude of both contributions is equal, so it depends on their phase relation what type of final FRF will develop⁴.

⁴ In Section 4.1 much attention is paid to the interference of modal contributions in FRF.

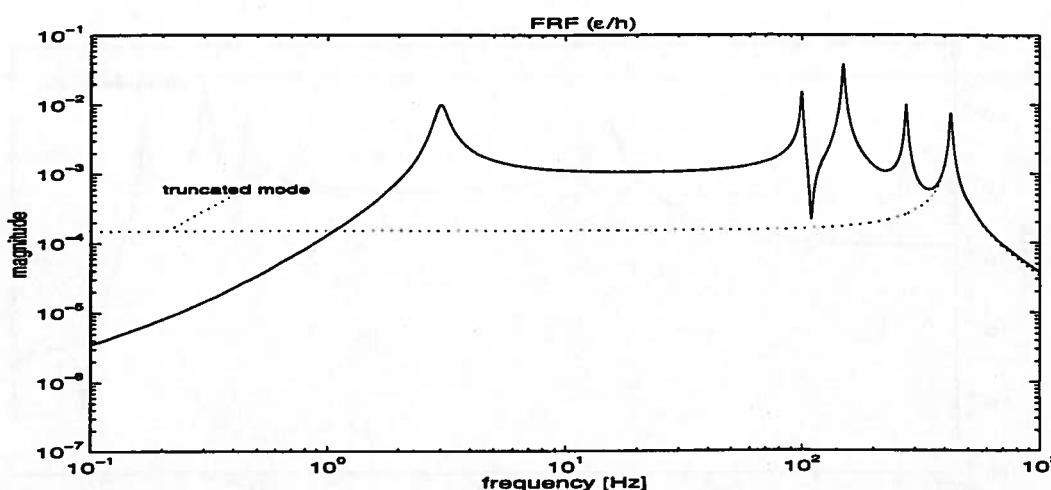


Figure 5.23 Sensitivity curve (ϵ/h) of complete model (solid : final FRF, dotted :single modal contributions)

Numerical accuracy

In general, as more modes are included in the analysis, the results become more in agreement with the expected results. However, even when all modes are included, one should not forget the effect of the finite precision with which numbers are represented in a computer. This limited numerical accuracy has a similar effect on the static solution as the truncation of modes.

To illustrate this effect, one can again consider the 2-DOF lumped model (Fig.5.18). The internal deformation ϵ results from a subtraction of the motion x_2 of the frame from the motion x_1 of the manipulator⁵, while each of these motions is the result of two modal contributions, which leads to :

$$\left(\frac{\epsilon}{h} \right) = \left(\frac{x_1 - x_2}{h} \right) = \left(\frac{x_1}{h} \right)_1 - \left(\frac{x_2}{h} \right)_1 + \left(\frac{x_1}{h} \right)_2 - \left(\frac{x_2}{h} \right)_2 \quad (5.9)$$

The low-frequency contribution of the first mode to each of the individual FRF $(x_1/h)_1$ and $(x_2/h)_1$ is almost equal to "1", whereas according to Fig.5.19 the final sensitivity (ϵ/h) at a frequency f ($f < f_1$) is approximately equal to :

$$\left(\frac{\epsilon}{h} \right)_{f < f_1} \approx \left(\frac{f}{f_1} \right)^2 * \left(\frac{f_1}{f_2} \right)^2 = \left(\frac{f}{f_2} \right)^2 \quad (5.10)$$

So, assuming a typical suspension frequency of 3Hz and an internal frequency of 100Hz, one finds a sensitivity of about 10^{-6} at 0.1Hz. This small number has to be calculated by subtracting modal contributions of order "1" from each other. In order not to obtain any unexpected results, the error of the mode-shape vectors needs to be at least 10-100 times smaller than this number.

⁵ From a numerical accuracy point of view this is unfavourable and it would be better to express the equations of motion directly in the variables x_2 and ϵ , but unfortunately this approach is generally not supported by the modelling software.

Final Remark

Despite the fact that very simple examples have been used to illustrate the effect of modal truncation and numerical accuracy, the phenomena have general validity and can be observed in the analysis of any detailed mechanical model which is based on a truncated modal description with limited numerical accuracy.

Probably the most important conclusion about modal truncation is that one should be aware of the deteriorating effect, that it can have on simulation results. Understanding the nature of these effects allows the analyst to take proper action. Either one can increase the number of included modes or the numerical accuracy, or one can simply try to live with the result. In the case of the floor vibration sensitivity, for example, it is perfectly legitimate - after having investigated the nature of the deviations from the expected results - to replace the low-frequency part of the simulated curve with a +2 slope up to the suspension frequency.

5.3.3. Model-Reduction Techniques

A complete coverage of the topic “model reduction” goes far beyond the scope of this thesis, but some comments need to be made on the suitability of popular approaches for the purpose of substructuring⁶:

- Static-Reduction Techniques
- Component-Mode Techniques

The most general type of component, or sub-structure, is one that is connected to one or more neighbouring components in a statically over-determined manner, as illustrated in Fig.5.24 by the example of a planar truss component. The example also shows the differentiation that is made between :

- boundary (or interface) DOF
- interior DOF

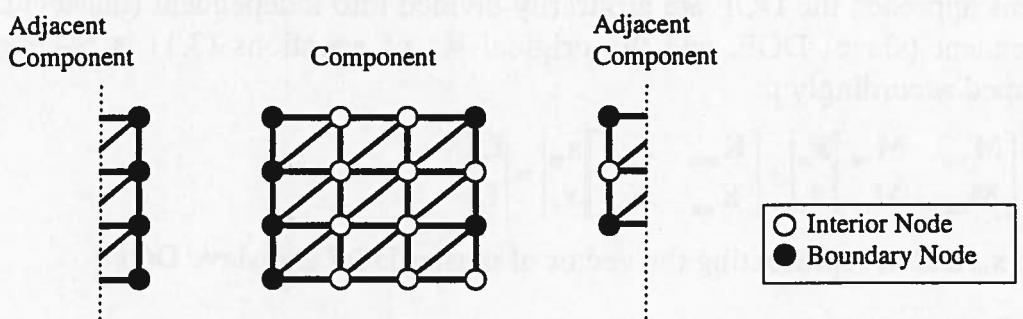


Figure 5.24 Example of a truss component that is connected to two neighbouring components

⁶ Model-reduction techniques which are developed for the purpose of model-based control design such as described in [WOR94] will not be discussed.

The fundamental goal of model-reduction techniques is to provide a reduced-size model while preserving the mechanical properties of interest. Especially the behaviour of the overall system model, consisting of several reduced component models, has to be correct. According to [CRA81], these techniques should have the following characteristics:

1. Accuracy of the reduced system model (system that is assembled from reduced components) up to a certain frequency needs to be ensured.
2. Efficiency, which means that the additional overhead, in terms of effort and time due to the reduction step of each individual component, must be regained in the analysis of the assembled system.
3. Component independence, such that one can model one component without knowing the dynamic properties of neighbouring components. This is especially important in case the modelling activity is distributed over several people, or one component is used in various assemblies⁷.
4. Synthesis generality implies that all types of components should be treated in one uniform way.
5. Static solution capability of components is a premise to ensure the performance of the assembled system under static loading conditions, but it is also important in view of the dynamic properties of the assembled system.
6. Explicit boundary co-ordinates avoid complicated coupling mechanisms based on generalised co-ordinates.

5.3.3.1. Static-Reduction Techniques

Static-reduction techniques, which have been developed in the past to enable the calculation of modes and natural frequencies of very large FE models, operate directly on the original mass and stiffness matrices, without prior calculation of any dynamic properties of the original model.

In this approach the DOF are arbitrarily divided into independent (master) DOF and dependent (slave) DOF, and the original set of equations (3.1) is re-ordered and grouped accordingly :

$$\begin{bmatrix} \mathbf{M}_{mm} & \mathbf{M}_{ms} \\ \mathbf{M}_{sm} & \mathbf{M}_{ss} \end{bmatrix} \begin{bmatrix} \ddot{\mathbf{x}}_m \\ \ddot{\mathbf{x}}_s \end{bmatrix} + \begin{bmatrix} \mathbf{K}_{mm} & \mathbf{K}_{ms} \\ \mathbf{K}_{sm} & \mathbf{K}_{ss} \end{bmatrix} \begin{bmatrix} \mathbf{x}_m \\ \mathbf{x}_s \end{bmatrix} = \begin{bmatrix} \mathbf{f}_m \\ \mathbf{f}_s \end{bmatrix} \quad (5.11)$$

with \mathbf{x}_m and \mathbf{x}_s representing the vector of master DOF and slave DOF.

⁷ In practice it is advisable to consider neighbouring components during the modelling process and to define the location of interface nodes, such that the coupling of sub-structures becomes significantly more straightforward.

If the deflections of the slave DOF \mathbf{x}_s are approximated by a linear combination of the deflections of the master DOF \mathbf{x}_m , the slave DOF \mathbf{x}_s can be eliminated from the original set of equations of motion (5.11). This leads to the following reduced approximation of the original set of equations :

$$\mathbf{M}^* \ddot{\mathbf{x}}_m + \mathbf{K}^* \mathbf{x}_m = \mathbf{f}^* \quad (5.12)$$

in which \mathbf{M}^* , \mathbf{K}^* and \mathbf{f}^* represent the new mass matrix, stiffness matrix and force vector.

More details about this approach can be found in Appendix B, which also shows a very illustrative example of the most prominent exponent of these techniques, namely the Guyan reduction.

The static-reduction technique can theoretically result in a significant reduction of the model size of a sub-structure, and the approach satisfies requirements 2 to 6 (providing that the boundary or interface DOF are included in the set of master DOF). However, this technique offers no means to guarantee the accuracy of the reduced model (requirement 1), despite the fact that certain general guidelines can be given, such as :

- the number of master DOF should be at least twice the number of modes of interest
- master DOF should be located at nodes with high relative inertia
- rotational DOF and DOF involving high relative stiffness (for example elongation of beam as compared to bending deformation) can in many cases be chosen as slave DOF
- in systems with distributed mass and stiffness (for example a homogeneous plate), the master DOF should also be distributed homogeneously.

The choice of master DOF is highly problem-dependent, and a badly chosen set of master DOF can seriously affect the dynamic properties of the component model. Moreover, there is no criterion for the required number and choice of master DOF in the component description, such that the results of the assembled system can be guaranteed up to a certain frequency. As a result of these limitations, the static-reduction techniques are considered less suitable.

5.3.3.2. Component-Mode Techniques

In the component-mode technique, the deflection of a sub-structure is approximated by a linear combination of m constant shape vectors (equation (5.13)), which are called component modes. These component modes are calculated on the basis of the original model.

This approach is very similar to the modal decomposition approach, which was discussed in Chapter 3, except for the fact that the component modes do not necessarily have to be mode-shapes as they are known from modal analysis. Various (static and dynamic) deflection shapes can be used as component modes (Appendix C illustrates some of the most often used component modes using the example of a simple planar beam structure). The collection of component modes ψ_i^8 ($i=1,2,\dots,m$) is called the component mode set. The displacement vector $x(t)$ is approximated by a linear combination of component modes ψ_i :

$$x(t) = p_1(t)\psi_1 + p_2(t)\psi_2 + \dots + p_m(t)\psi_m = \Psi p(t) \quad (5.13)$$

with

Ψ : $n*m$ matrix of m preselected component modes with n DOF each

$p(t)$: vector with m generalised co-ordinates

Substitution of equation (5.13) in the original equation of motion (3.1), and premultiplication with Ψ^T , yields the following description, which is an approximation on the basis of energy equivalence :

$$M_c \ddot{p} + K_c p = \Psi^T f \quad (5.14)$$

with

$M_c = \Psi^T M \Psi$: reduced mass matrix ($m*m$)

$K_c = \Psi^T K \Psi$: reduced stiffness matrix ($m*m$)

This transformation reduces the size of the problem to m equations in the generalised DOF p . In contrast to the modal mass and stiffness matrix, the reduced matrices M_c and K_c in general do not have a diagonal structure.

It is of utmost importance to properly choose the component modes such that the dynamic properties are well approximated. In literature various component modes are found. These component modes can be combined in different ways, leading to a large variety of component-mode sets. On the basis of a literature scan in combination with common practice and availability of supporting software, the use of two component-mode sets has been further evaluated [VEL89], namely :

- free-interface normal modes⁹
- Craig-Bampton mode set

A brief summary of the main results of this investigation will be given here.

⁸ In order to avoid any confusion with the mode-shape vector ϕ_i , a different symbol has been introduced to indicate a component mode, namely ψ_i .

⁹ Normal modes are the eigenvectors which result from the solution of the eigenvalue problem for the unforced vibration of the substructure. Depending on the type of boundary conditions, one can distinguish between free-interface, fixed-interface, hybrid-interface, and loaded-interface modes.

Free interface normal modes

Sub-structure reduction based on free-interface normal modes is fairly common. In this approach each component of the system is described by a limited number of mode-shapes and modal parameters under free-free support (or interface) conditions.

This approach satisfies the requirement for efficiency (2), component independence (3) and synthesis generality (4). The set of generalised co-ordinates does not include the boundary DOF in an explicit manner (6) and therefore more elaborated coupling procedures are required. However, the most important disadvantage of this technique is the lack of static solution capability (5), as already illustrated in Section 5.3.2. This not only affects the static solution of the assembled system, but also the accuracy (1) of the modal results of the assembled system.

This is not surprising, because the lower modes describe the global behaviour of a component, whereas the local flexibility near an interface DOF is described by the higher modes, which have been truncated. Consequently, the quality of the dynamic results of the assembled system can be seriously affected.

The quality of the component description can dramatically be improved by the addition of a so-called “residual flexibility matrix”, which accounts for the flexibility of the deleted modes and thus ensures the static solution capability. This method has not been further evaluated, as it has no benefits over the Craig-Bampton approach, which is discussed next.

Craig-Bampton approach

In 1968 Craig and Bampton introduced a component mode set [CRA68], which consists of :

- Constraint modes¹⁰
- Fixed-interface normal modes

A so-called Craig-Bampton component is defined by all its constraint modes (the number of these static deflection shapes corresponds to the number of interface DOF of the component) plus a limited number of mode-shapes and modal parameters under fixed interface conditions. It can be shown that this approach satisfies all the requirements, which were specified earlier.

In the investigation it was found that the required frequency range of the system solution is a good indication for the highest fixed-interface normal mode that must be included in each component description to guarantee the quality of the system result. A more conservative indicator has been reported by [WAN81] who claims

¹⁰ Constraint modes are static deflection shapes which are obtained by successively imposing a unit deflection in one of the boundary DOF while keeping the other boundary DOF fixed.

that the highest fixed-interface mode frequency should be about 35% higher than the highest frequency one needs to consider in the analysis of the assembled system.

Due to the fact that the Craig-Bampton approach satisfies all requirements of a good sub-structuring technique, and especially that it can guarantee the frequency up to which the results of the assembled system are reliable, makes it by far superior to the static-reduction technique and the component-mode technique based on free-interface normal modes.

6. Case Study of a CD-drive

In the beginning of the eighties, the well-known compact disc drive entered the market place and gradually took over from the traditional mechanical pick up.

Designing a unit that under all environmental conditions (for example -40 to +80° temperature), user conditions (for example portable systems used during jogging), and production tolerances, achieves sub-micron accuracy at a discount price, is a major challenge for the entire design team. The control of the dynamics and its influence on the performance of the servo system is one of the important aspects, and will probably remain an important issue in the future, because specifications in terms of critical dimensions and speed are ever increasing.

Until about 1985 the detrimental effects of mechanical vibrations have been tackled by a very time-consuming process of "trial-and-error". After finishing a design and assembling the first prototypes, experiments were done, and only then could the designs be experimentally evaluated from dynamics and control points of view. On the basis of these measurements and engineering judgement, changes in the mechanical design were made, followed by another time-consuming cycle of fabrication, assembly, and test of the new prototype.

In that time various groups within Philips started using advanced modelling tools in order to better understand, and hopefully predict, the performance of compact disc modules. The result of these initial attempts was very limited because the industrial application of these tools proved to be complex, time-consuming, and disappointing. Results were late, incomplete or in the wrong "format", and consequently contributed only very little to the actual design process. The main reason for the failure of these individual attempts was the lack of "critical mass" to build up sufficient knowledge about the application area and modelling issues, and to develop operational skills of the modelling tools. Because of this, a project team was created in 1987 to build up the required expertise in the field of predictive modelling of dynamics and control related to optical disc drives. In this team, participants from various development, research, and support departments were combined. After a co-operation of more than two years, during which an enormous amount of insight and operational knowledge was gained, and which really boosted the field of predictive modelling of mechatronic positioning devices, the project team gave its final presentation in 1989 [EIJ89/2].

One of the carriers of the project team was a modular concept called CDM-8. This carrier, in which the author was intensively involved, will serve as an industrial case to illustrate the use of modelling tools in the design of a mechatronic positioning device. After a short introduction of the basics of a compact disc system, the concept of the CDM-8 drive, its specifications, and the various phases in the modelling process will be described and illustrated with representative simulation results.

6.1. Basics of a Compact Disc System

A compact disc module (Fig.6.1) basically consists of a turntable with a well-regulated rotational speed on which the compact disc is placed, and an optical pick-up unit that reads the information from the surface of the disc. The information is digitally coded and engraved in a spiral track of pits.

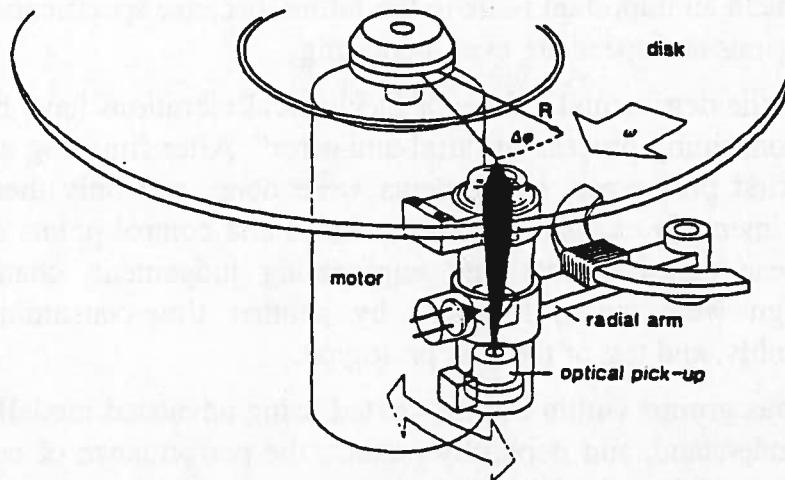


Figure 6.1 Schematic representation of a single-beam swing-arm compact disc module

The optical pick-up unit of a compact disc module consists of a laser, a semi-transparent mirror, a lens, and an array of four photo-diodes. The light beam is emitted by the laser, passes through a series of optical devices, is reflected by the disc, passes again through a series of optical components, and is then led to the photo-diodes. The digital information on the disc is transformed into intensity variations of the reflected laser light. Pits on the disc absorb the laser light (via interference), whereas the rest of the disc reflects the light. Via optical means and processing of the distribution of the reflected laser light over the four photo-diodes, it is furthermore possible to extract the focus and radial tracking errors [DOR86].

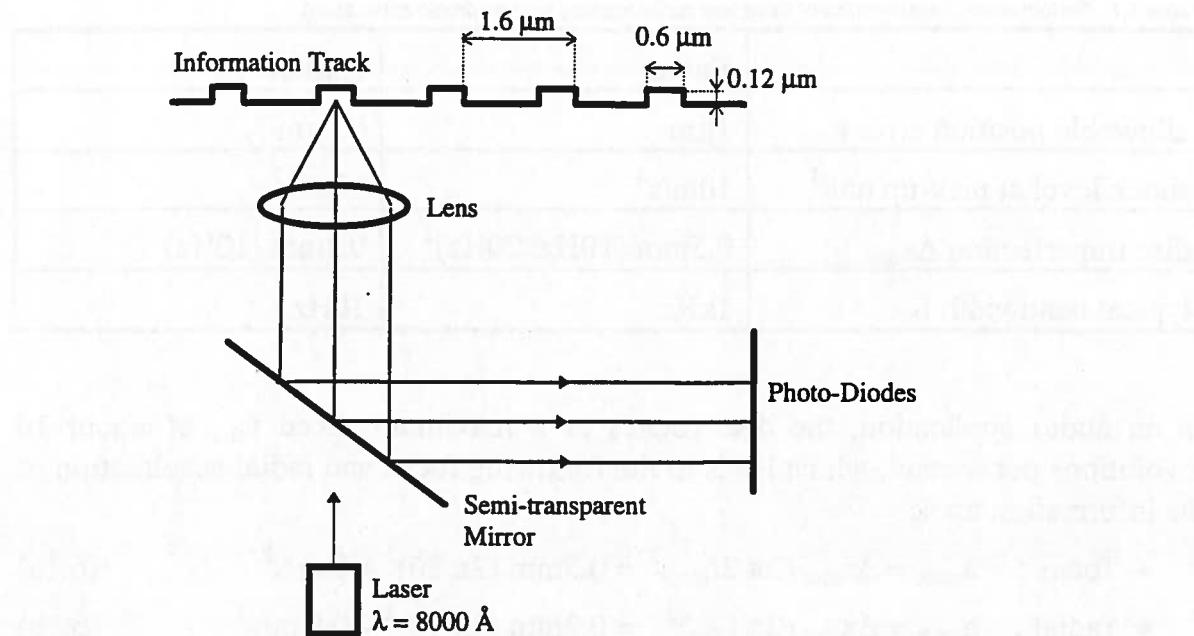


Figure 6.2 Path of emitted laser light from laser to array of photo-diodes

In audio application, the information on the disc is read with a constant linear speed of 1.25m/s , which corresponds to a rotational speed of the disc that varies from 3 to about 10 revolutions per second. Modern CD-ROM players are operated at speeds up to 16 times higher.

6.2. Servo Specifications

In the previous section the basic characteristics of a compact disc drive were discussed. To derive performance specifications for the focus and radial control-system, the functional requirements regarding positional accuracy and the disturbance mechanisms need to be investigated.

External shocks imposed on a CD player, which are transferred via the suspension mounts to the optical pick-up unit, are one source of disturbances. The other source is formed by the imperfection of the disc itself, which introduces radial and vertical motions of the track. Eccentricity of the information track and the axis of rotation of the disc results in radial motion of the track, whereas unflatness of the disc leads to vertical displacements of the track.

Table 6.1 gives an overview of the tolerable positioning errors, shock levels at the optical pick-up unit due to external shocks applied to the CD-player, disc imperfections, and the required servo performances in terms of bandwidth. The latter can be derived by considering the focus and radial accelerations of the track which have to be “followed” by the actuator.

Table 6.1 Performance characteristics of focus and radial tracking system (audio application)

	Focus	Radial
allowable position error ϵ	1 μm	0.2 μm
shock level at pick-up unit ¹	10 m/s ²	10 m/s ²
disc imperfection Δx_{disc}	0.5 mm (10Hz, 20Hz) ²	0.2 mm (10Hz)
typical bandwidth f_b	1 kHz	1 kHz

In an audio application, the disc rotates at a maximum speed f_{disc} of about 10 revolutions per second, which leads to the following focus and radial acceleration of the information track :

$$\bullet \text{ focus : } a_{\text{track}} = \Delta x_{\text{disc}} (2\pi 2f_{\text{disc}})^2 = 0.5 \text{ mm} (2\pi 20)^2 \approx 8 \text{ m/s}^2 \quad (6.1a)$$

$$\bullet \text{ radial : } a_{\text{track}} = \Delta x_{\text{disc}} (2\pi f_{\text{disc}})^2 = 0.2 \text{ mm} (2\pi 10)^2 \approx 0.8 \text{ m/s}^2 \quad (6.1b)$$

These acceleration levels are lower than those resulting from external shocks ($\approx 10 \text{ m/s}^2$) and are therefore less important in an audio application³.

For a first estimate of the required bandwidth one could neglect the influence of the integrator action of the controller and investigate the FRF ($\epsilon/a_{\text{track}}$) of a controlled system with bandwidth ω_b , which is shown in Fig.6.3. The low-frequency asymptote equals $1/\omega_b^2$, which leads to the following bandwidth requirement in radial direction

$$\left(\frac{\epsilon}{a_{\text{track}}} \right)_{\text{low-frequency}} = \frac{1}{\omega_b^2} = \frac{1}{(2\pi f_b)^2} < \left(\frac{0.2 \mu\text{m}}{10 \text{ m/s}^2} \right) \Rightarrow f_b > 1 \text{ kHz} \quad (6.2)$$

Because of the larger tolerable position error in focus direction, the bandwidth of the focus servo-system could be chosen at about 500Hz. Nevertheless it is generally also set at 1kHz because of vertical track accelerations which sometimes occur as a result of higher order disc unflatness.

This initial estimate of the bandwidth can be used as a starting-point in the dynamic analysis of the compact disc module. It was shown in Section 4.2 that the destabilising effect of a mechanical resonance depends on the frequency-range in which it occurs. On the basis of default controller-settings (Fig.4.7), the results shown in Fig.4.8, and a bandwidth of 1kHz the three frequency-ranges can be defined as :

¹ The suspension mounts of the pick-up unit are designed such that external shocks imposed on the CD-player (50 m/s^2 , 2-3ms) are reduced to a level of about 10 m/s^2 at the pick-up unit.

² A saddle-shaped unflatness of the disc leads to a sinusoidal vertical displacement of the track with twice the rotational frequency of the disc.

³ Increasing the rotational speed of the disc significantly, as in modern CD-ROM modules, makes the accelerations resulting from disc imperfections become dominant.

- low-frequency : < 200Hz
- medium-frequency : 200-3500Hz
- high-frequency : > 3500Hz

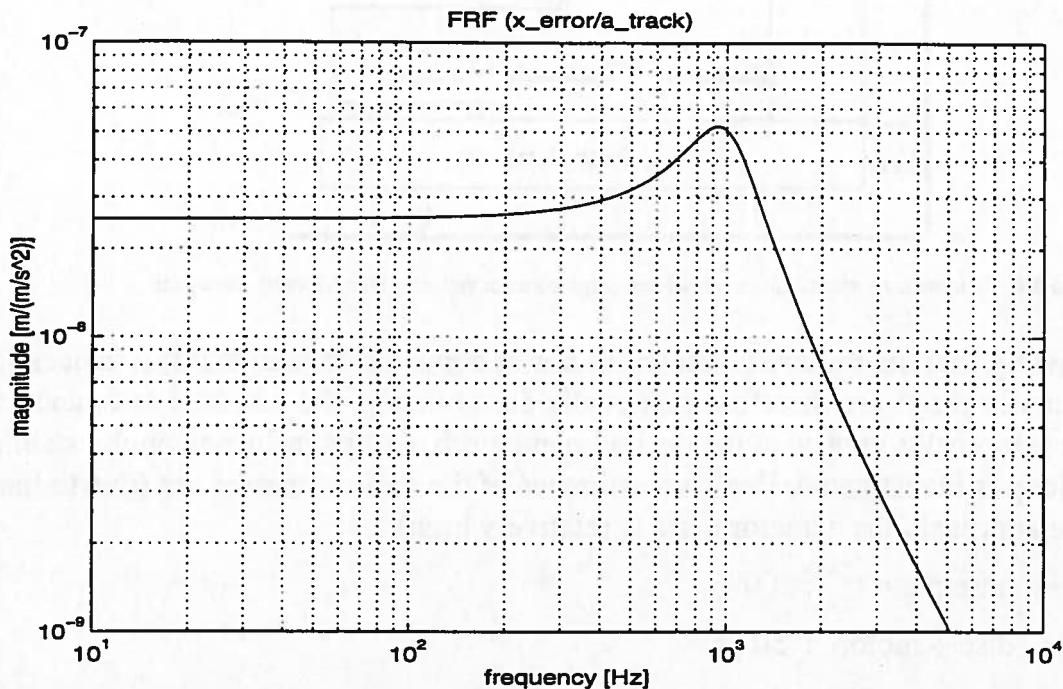


Figure 6.3 FRF (ϵ/a_{track}) in the absence of integral action

6.3. Concept Evaluation

The comparison of a one-stage radial concept versus a two-stage radial concept regarding the stability of the control loop is an interesting example of a conceptual evaluation.

Figure 6.4. shows a schematic representation of an optical pick-up unit based on a one-stage tracking configuration (although linear motions are considered here, such a model can also be derived for a rotational tracking system as shown in Fig.6.1). The radial tracking is done by one single device that is used for both the large motion (inner to outer track) as well as for small correcting movements to compensate for disc eccentricity or disc motion due to shocks. In this first conceptual model, the limited mass and supporting stiffness of the base plus one resonance of the disc and disc motor with respect to the base plate, are considered.

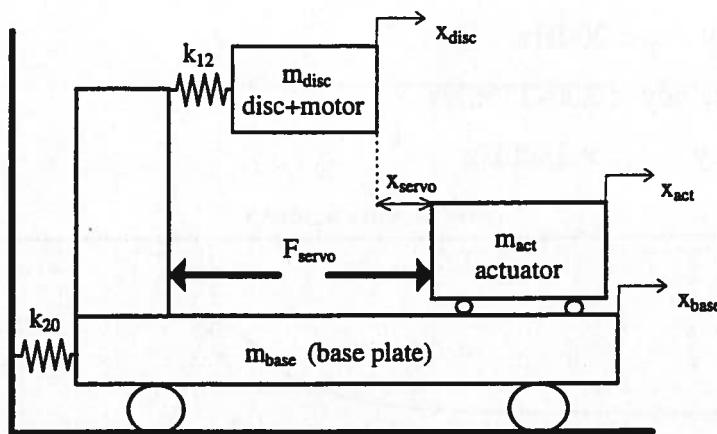


Figure 6.4 Schematic representation of radial one-stage concept including disc (+motor) dynamics

Knowing that the required bandwidth lies at about 1kHz, and that resonances in that frequency range are therefore potentially destabilising, the parasitic resonance of the disc (+motor) is chosen at the desired bandwidth, and its influence on the stability of the loop is investigated. Realistic estimates of the various masses are (due to the one-stage approach, the actuator mass is relatively high):

- base plate : 200 gr
- disc + motor : 50 gr
- actuator : 20 gr

The suspension frequency, which typically lies at 25Hz, has a relative damping of 5%, whereas the internal mode at 1kHz is assumed to have only 1% relative damping. The resulting FRF ($x_{\text{servo}}/F_{\text{servo}}$) is shown in Fig.6.5.

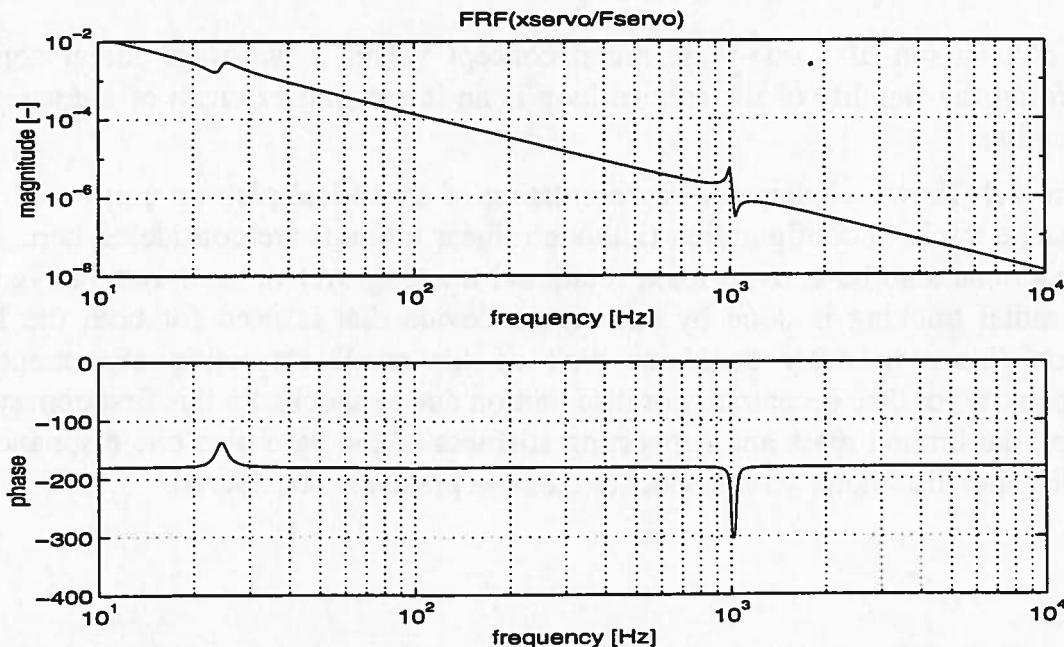


Figure 6.5 FRF ($x_{\text{servo}}/F_{\text{servo}}$) corresponding to the system shown in Fig.6.4.

The well-damped resonance at 25Hz poses no stability problem. However, the resulting characteristic of the disc (+motor) resonance at 1kHz is of the type -2slope/pole/zero/-2slope, which near the desired bandwidth frequency seriously endangers the stability of the loop.

It is interesting to investigate how this FRF can be improved such that the influence of this resonance decreases. In order to do so it is essential to understand how this FRF is constructed from the various modal contributions, and how the various variables influence its shape. An examination of the mode-shapes is always a good starting point for such an investigation. This can be done without any sophisticated mathematics if one considers that the natural frequency of the suspension mode (25Hz) and that of the internal mode (1kHz) are so far apart that at 25Hz the disc (+motor) are almost rigidly connected to the base plate, whereas at 1kHz the suspension stiffness of the base plate can be neglected.

These considerations lead to the mode shapes, which are listed in Table 6.2 and graphically visualised in Fig.6.6.

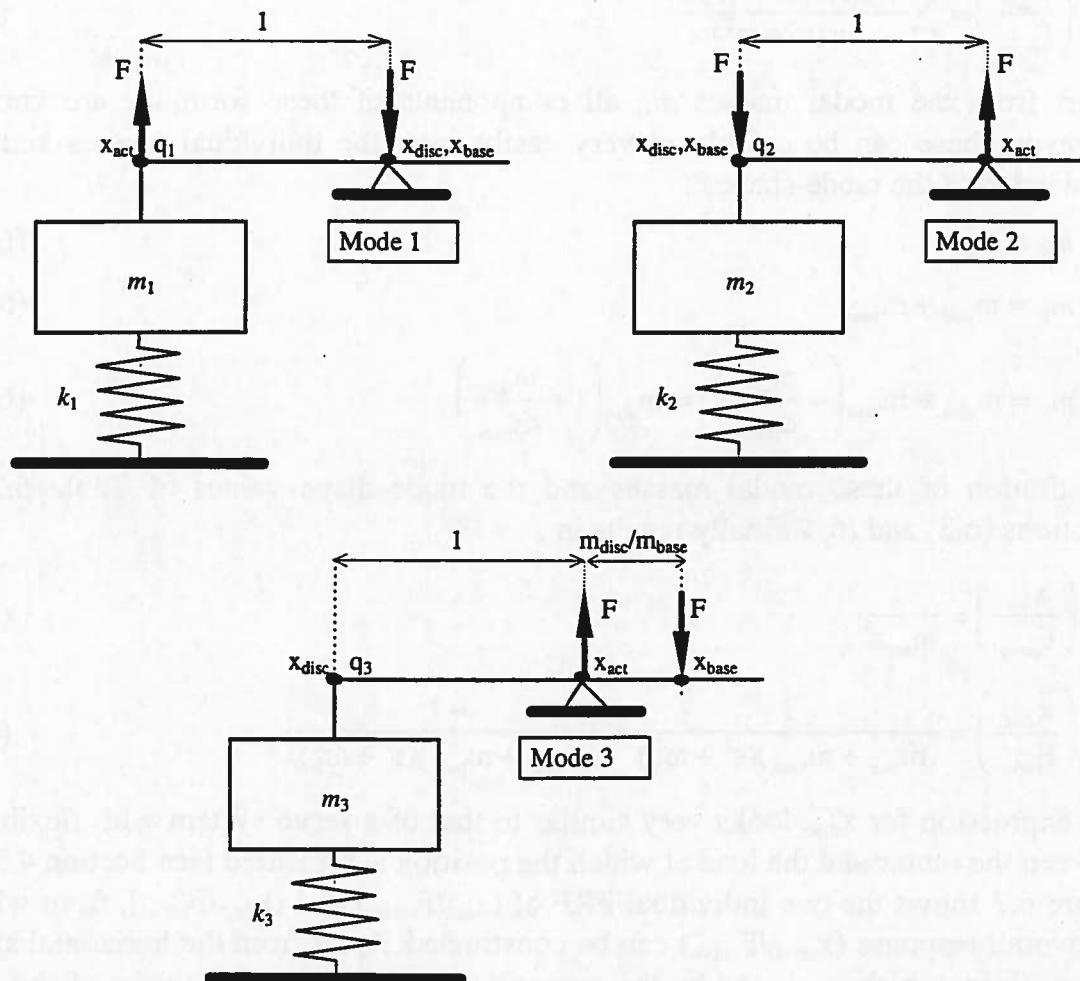


Figure 6.6 Graphical representation of mode-shapes listed in Table 6.2.

Table 6.2 Approximated mode-shapes corresponding to the system of Fig.6.4.

	Mode 1 (0 Hz)	Mode 2 (25Hz)	Mode 3 (1kHz)
$\phi_{i,act}$	1	0	0
$\phi_{i,disc}$	0	1	1
$\phi_{i,base}$	0	1	$-m_{disc}/m_{base}$

Once these modes are known, the individual FRF (x_{act}/F_{servo}) and (x_{disc}/F_{servo}), from which the overall FRF (x_{servo}/F_{servo}) is derived, can be calculated. When doing so, one has to realize that the servo force F_{servo} , which acts on the actuator, also has a reaction component which acts in the opposite direction on the base. Considering the action and reaction servo force, the formulae for the two FRF read :

$$\left(\frac{x_{act}}{F_{servo}} \right) = \sum_{i=1}^3 \frac{\phi_{i,act} (\phi_{i,act} - \phi_{i,base})}{m_i (s^2 + \omega_i^2)} \quad (6.3)$$

$$\left(\frac{x_{disc}}{F_{servo}} \right) = \sum_{i=1}^3 \frac{\phi_{i,disc} (\phi_{i,act} - \phi_{i,base})}{m_i (s^2 + \omega_i^2)} \quad (6.4)$$

Apart from the modal masses m_i , all components of these formulae are known; however, these can be calculated very easily from the individual masses and the knowledge of the mode-shapes :

$$m_1 = m_{act} \quad (6.5a)$$

$$m_2 = m_{disc} + m_{base} \quad (6.5b)$$

$$m_3 = m_{disc} + m_{base} \left(-\frac{m_{disc}}{m_{base}} \right)^2 = m_{disc} \left(1 + \frac{m_{disc}}{m_{base}} \right) \quad (6.5c)$$

Substitution of these modal masses and the mode-shape values of Table 6.2 in equations (6.3) and (6.4) finally results in :

$$\left(\frac{x_{act}}{F_{servo}} \right) = \frac{1}{m_{act} s^2} \quad (6.6)$$

$$\left(\frac{x_{disc}}{F_{servo}} \right) = \frac{1}{(m_{disc} + m_{base})(s^2 + \omega_2^2)} + \frac{-1}{(m_{disc} + m_{base})(s^2 + \omega_3^2)} \quad (6.7)$$

The expression for x_{disc} looks very similar to that of a servo system with flexibility between the motor and the load at which the position is measured (see Section 4.3.1). Figure 6.7 shows the two individual FRF of (x_{act}/F_{servo}) and (x_{disc}/F_{servo}), from which the overall response (x_{servo}/F_{servo}) can be constructed. Apart from the horizontal slope below 25Hz, which is caused by the suspension stiffness, the behaviour of the disc

indeed shows the expected -2slope/pole/-4slope combination at 1kHz. The asymptote of $(x_{\text{disc}}/F_{\text{servo}})$ between ω_2 and ω_3 is defined by :

$$\left(\frac{x_{\text{disc}}}{F_{\text{servo}}} \right)_{\omega_2^2 < s^2 < \omega_3^2} = \frac{1}{(m_{\text{disc}} + m_{\text{base}})s^2} \quad (6.8)$$

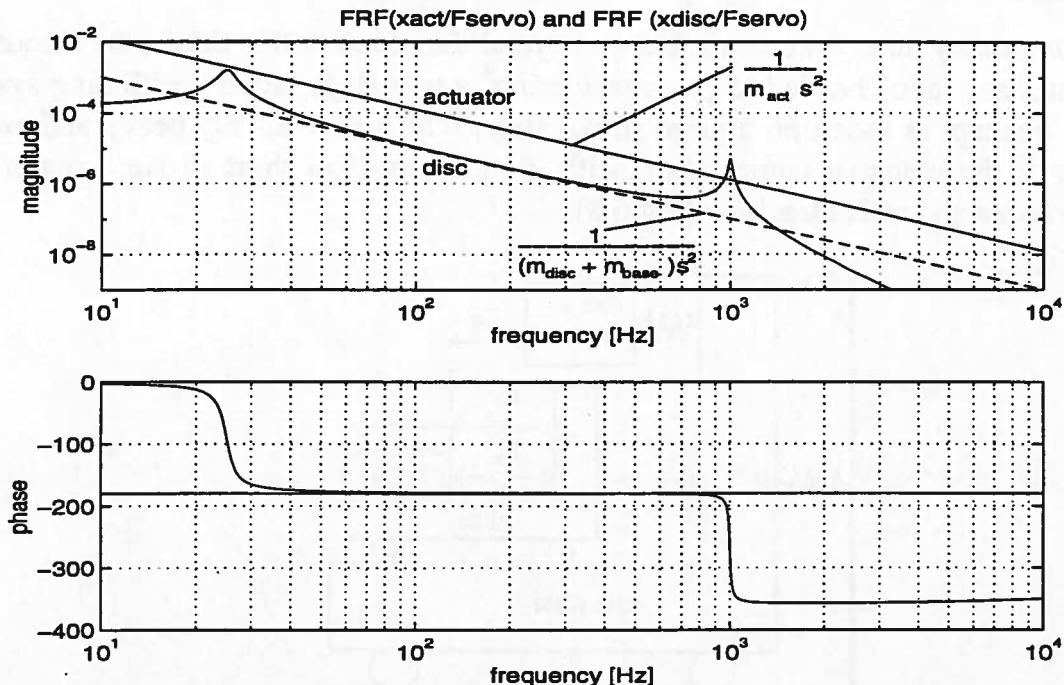


Figure 6.7 Individual FRF of $(x_{\text{act}}/F_{\text{servo}})$ and $(x_{\text{disc}}/F_{\text{servo}})$

Figure 6.7 also shows that the amount of interference between the desired motion of the actuator and the motion of the disc (and consequently the amount of destabilisation due to the 1kHz resonance of the disc) is completely determined by the mass ratio $m_{\text{act}}/(m_{\text{base}}+m_{\text{disc}})$ and the amount of resonance amplification Q of the 1kHz mode. A good measure for the relative importance of the parasitic mode compared to the desired motion is given by :

$$Q \frac{m_{\text{act}}}{m_{\text{base}} + m_{\text{disc}}} \quad (6.9)$$

Substituting the values of the example into equation (6.9) leads to :

$$Q \frac{m_{\text{act}}}{m_{\text{base}} + m_{\text{disc}}} = 50 \frac{20}{200+50} = 4 \quad (6.10)$$

which indicates that at the resonance frequency the corresponding peak of the parasitic mode is four times higher than the desired motion. It can be shown that in order to reduce the resulting phase lag to an acceptable level of, for example 5°, the ratio $Q*m_{\text{act}}/(m_{\text{base}}+m_{\text{disc}})$ should be less than 0.1.

Increasing the mass of the base plate could help to reduce the effect of the parasitic mode, however this approach is commercially not feasible due to the miniaturisation and portability trend in consumer products. The room for a significant reduction of the actuator mass is also limited in this one-stage concept, because of the large stroke that needs to be travelled and because the mass of the components attached to the actuator puts a serious threshold to the mass reduction efforts⁴. Nor is a rigorous increase of the internal damping considered realisable.

Among many other arguments that go beyond the scope of this thesis, the discussion about mass ratios has led to (re-)considering⁵ a two-stage radial positioning system. This concept is based on a large stroke slide, which is relatively heavy and carries most of the essential components, with, on top, an extra short stroke actuator that does focus and radial tracking (Fig.6.8).

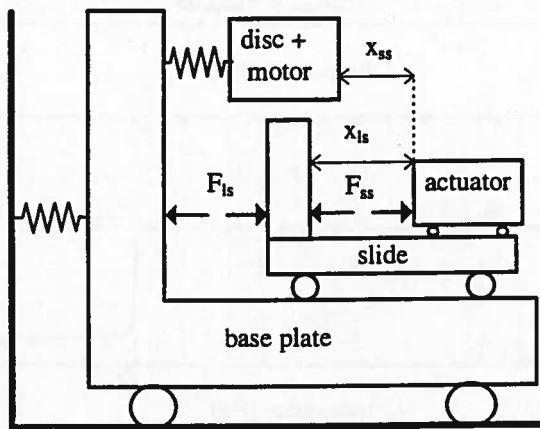


Figure 6.8 Two-stage tracking system

The small actuator is driven by Lorentz forces generated between the actuator and the large stroke slide, while the reflected laser spot is used as position information for the high frequency second stage servo loop. The control loop of the large stroke slide does not require a high bandwidth (typically 10-50Hz) because its only task is to follow the average position of the second stage such that the latter remains within its operating range ("stationary" part and moving part of the second stage should not deviate too much, due to the limited stroke of the actuator). This tolerable motion, which is related to the operating stroke of the actuator (typically +/-0.2mm), is orders of magnitudes larger than the tolerable motion of the short stroke actuator relative to the information track on the disc. Consequently, the requirements regarding the bandwidth of the slide servo are much more relaxed.

⁴ In spite of the suggestion made in the previous discussion, the one-stage concept can and does perform according to specifications, but due to the large excitation level of the mounting plate it requires very careful design of practically all components in order to guarantee servo stability.

⁵ The idea of a two-stage concept is not new. The predecessor of the compact disc, the Video Disc, was already equipped with a two-stage concept.

Various concepts can be used for the drive system, position-sensing, and control configuration of the large stroke actuator. One of the possible concepts comprises a linear direct-drive Lorentz actuator between the base plate and the slide, in combination with a position loop in which the radial position of the 2D-actuator relative to the slide is measured and controlled to a zero reference (Fig.6.9).

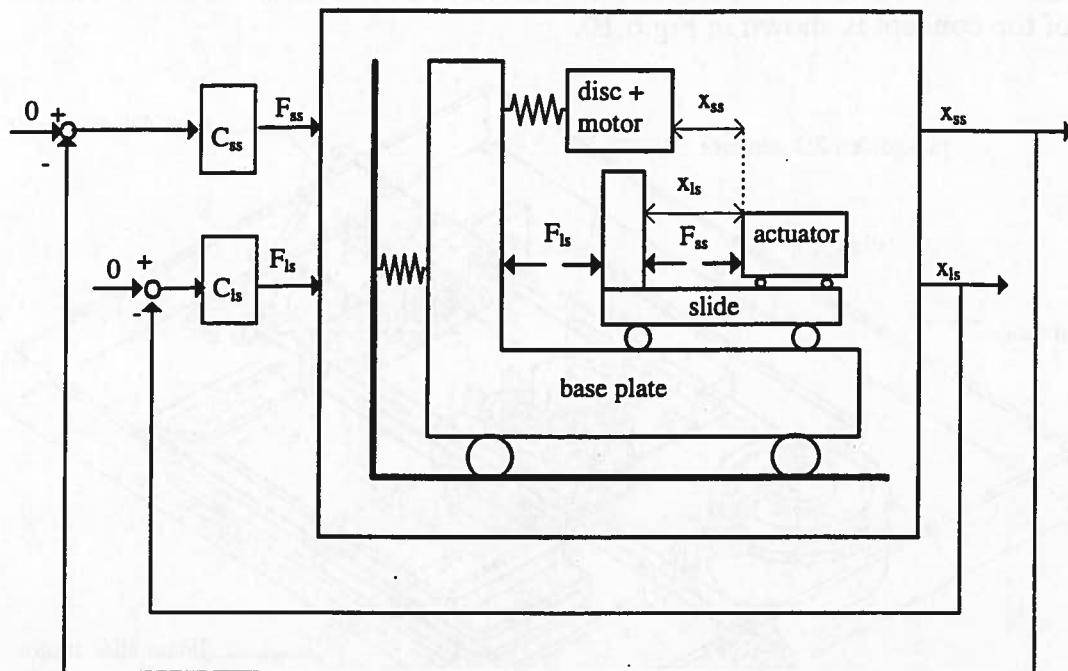


Figure 6.9 Possible concept for a two stage radial tracking system

Numerical evaluation of this radial two-stage concept, based on the model shown in Fig.6.9 has proven its viability. Due to its low servo-bandwidth specifications, the slide control poses no problem at all. The radial control-loop of the small stroke actuator shows the expected improvement compared to the original one-stage concept which is caused by the reduced mass of the actuator (1.25gr instead of 20gr). Due to the large ratio between moving and stationary masses, the mode of the disc (+motor) with respect to the base plate is practically invisible in the FRF of the small stroke actuator. In focus direction no significant change of the performance is expected because the concept of the focus servo system is still very similar to that of a single-stage swing-arm concept.

6.4. System Evaluation

On the basis of many considerations, including the previous discussion on mass ratios and servo stability, the choice was made to develop a modular concept based on a linear two-stage approach. This modular concept was called CDM-8 (Compact Disc Module) and the first application, a CD-ROM, was called CDM-8.1. The basic idea of the concept is shown in Fig.6.10.

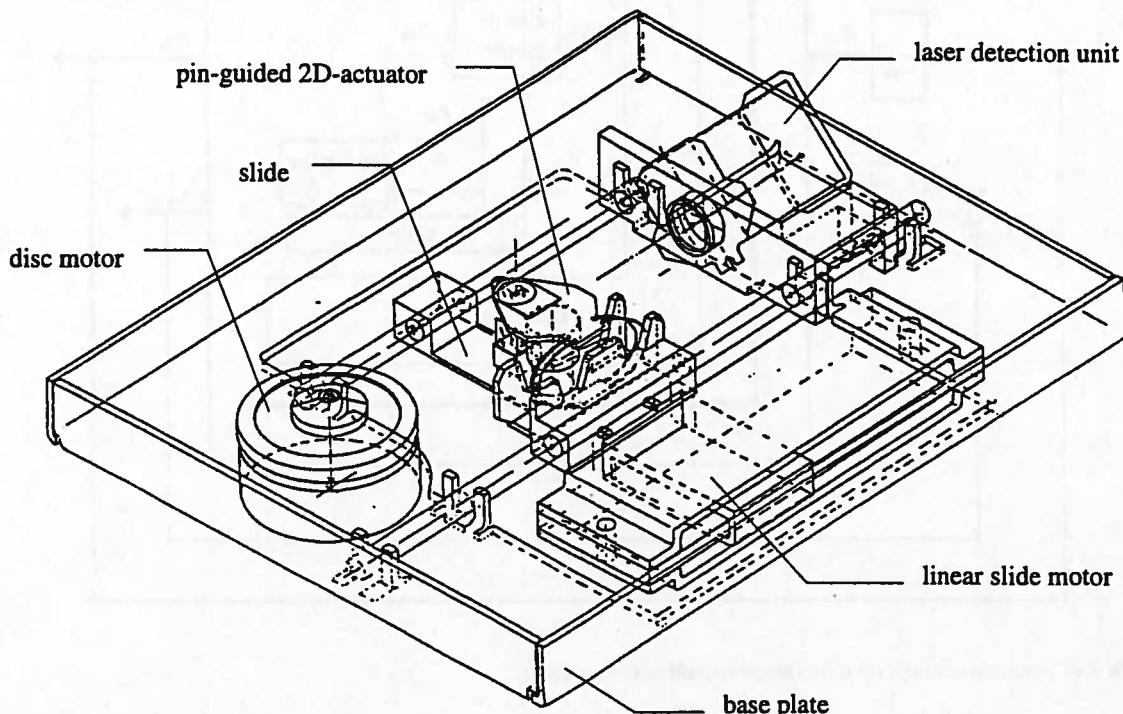


Figure 6.10 Prototype version of CDM-8.1

The system consists of a base plate which is connected with flexible rubbers to the CD player. On the base plate a disc motor with turntable that holds the disc and clamping device is assembled. A slide, driven by a linear Lorentz motor makes the large radial movement from the inner to the outer track of the disc, whereas a light-weight pin-guided 2D-actuator is used to “follow” the radial and focus motions of the information track that result from shocks and geometric imperfections of the disc. In this concept a laser detection unit, which is connected to the base plate, is applied to create a well-defined laser beam and process the reflected beam. As indicated in Fig.6.11, the light emitted by the laser is deflected by a corner mirror mounted on the slide, and travels through the lens which is part of the 2D-actuator.

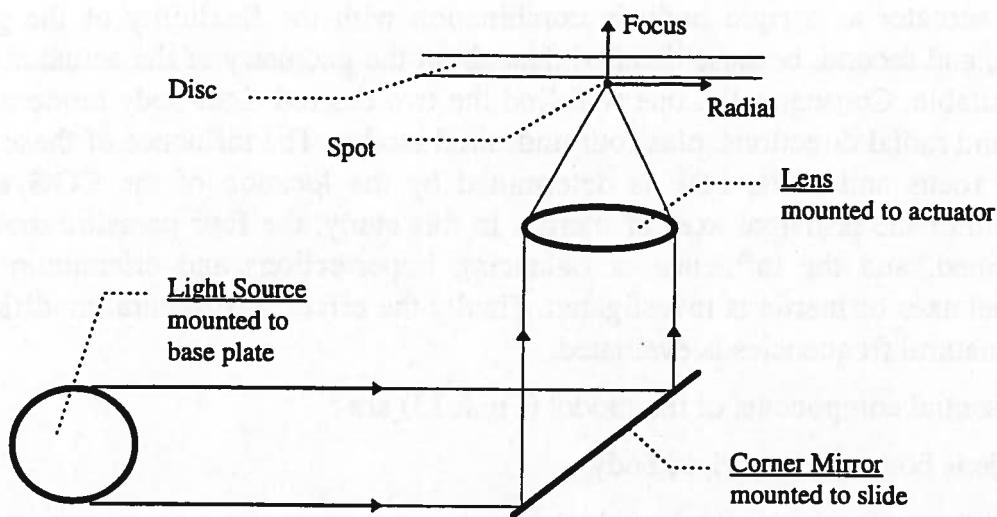


Figure 6.11 Laser light path of CDM-8 concept

6.4.1. Preliminary Actuator Model

On the basis of previous actuator designs, the choice was made to develop a pin-guided actuator⁶ (Fig.6.12) with two degrees of freedom: a linear focus motion (f), and a rotation about the guiding pin that leads to a radial displacement (r) of the lens (perpendicular to the track on the disc)⁷. The remaining four DOF are restrained by the guiding pin in combination with two jewelled slide-bearings.

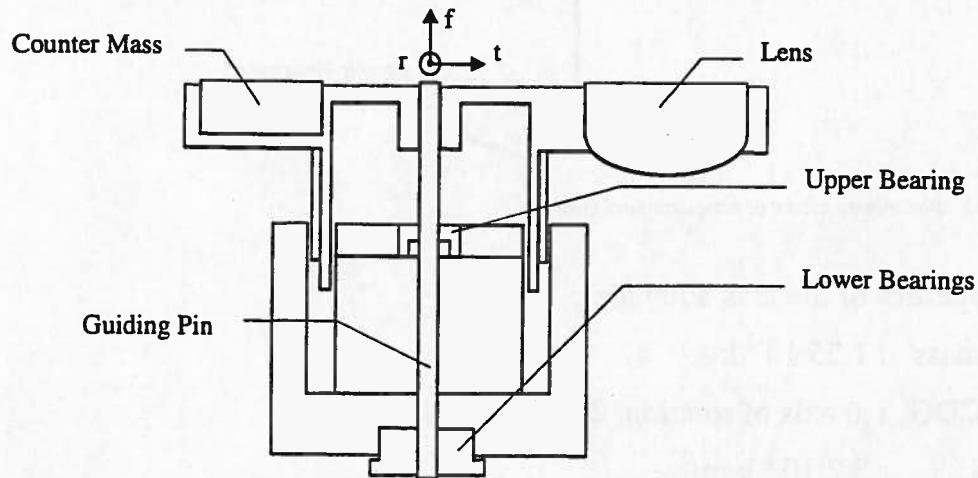


Figure 6.12 Pin-guided actuator

In the preliminary analysis [RAN88/1], the influence of the internal modes of the actuator body, which in view of the required bandwidth are targeted at 10kHz or higher, is not included. First, because the analysis focuses on the dynamic behaviour

⁶ This is a completely different type of actuator design than the one shown in Section 4.3.4.

⁷ The direction that corresponds to the track on the disc is indicated by "t" (tangential).

of the actuator as a rigid body in combination with the flexibility of the guiding system, and second, because detailed data about the geometry of the actuator are not yet available. Consequently, one will find the two desired rigid-body motions in the focus and radial directions, plus four undesired modes. The influence of these modes on the focus and radial FRF is determined by the location of the COG and the direction of the principal axes of inertia. In this study, the four parasitic modes are determined, and the influence of balancing imperfections and orientation of the principal axes of inertia is investigated. Finally the effect of structural modifications on the natural frequencies is evaluated.

The essential components of the model (Fig.6.13) are :

- lens body : rigid body
- pin : 2 beam elements
- guiding stiffness : discrete springs

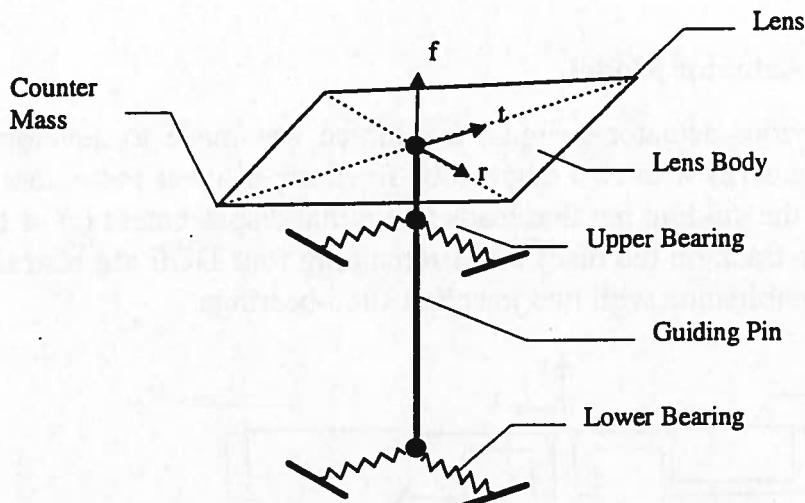


Figure 6.13 Preliminary model of pin-guided actuator

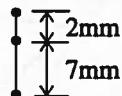
The properties of the lens body are :

- mass : $1.25 \cdot 10^{-3}$ kg
- COG : at axis of rotation, 2.7mm below top
- J_{rr} : $3.2 \cdot 10^{-8}$ kgm 2
- J_{tt} : $1.6 \cdot 10^{-8}$ kgm 2
- J_{ff} : $4.5 \cdot 10^{-8}$ kgm 2

The mass moments of inertia J_{rr} and J_{ff} are zero, because the design is symmetric with respect to the tf-plane. Initially J_{ff} is also set to zero; however, as this is not guaranteed without special design attention, the influence of this term is investigated.

The guiding pin is modelled by two beam elements (the minimum number of elements to create the three nodes which are required for the interface to the lens body and the two guiding locations). The properties of the guiding pin are :

- diameter d : 0.6 mm
- material : steel

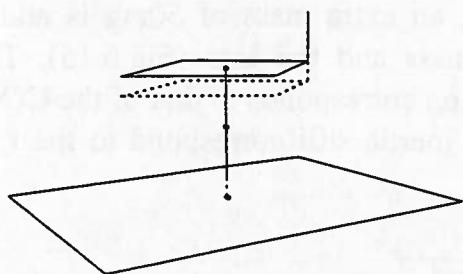


The support stiffness of the guiding points is initially set to 10^7 N/m .

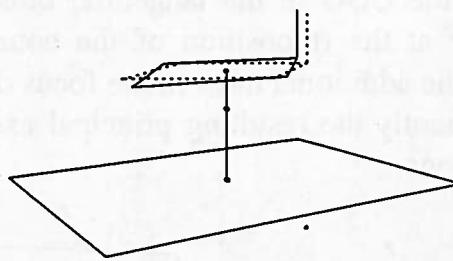
Reference model

First, the natural frequencies and modes (Fig.6.14) of the balanced reference model with its principle axes in r, t, and f directions ($J_{tf}=0$) are determined :

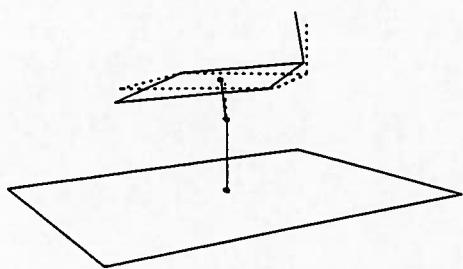
- 0 Hz : translation in focus direction (rigid body mode)
- 0 Hz : rotation about focus axis (rigid body mode)
- 490 Hz : rotation about radial axis (pitch mode)
- 690 Hz : rotation about tangential axis (roll mode)
- 3630 Hz : translation in tangential direction
- 3640 Hz : translation in radial direction



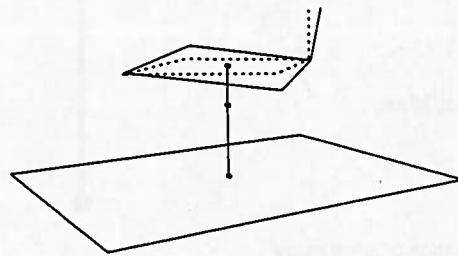
mode 1 : focus translation (0 Hz)



mode 2 : rotation about focus axis (0 Hz)



mode 3 : pitch (490 Hz)



mode 4 : roll (690 Hz)

(continued on next page)

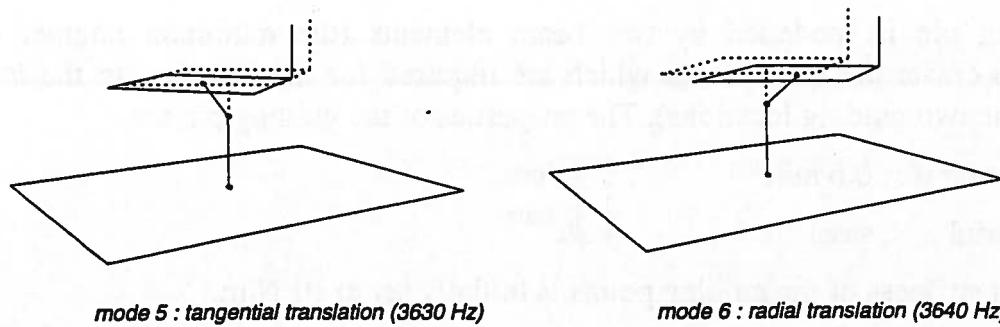


Figure 6.14 Modes of reference model (balanced, $J_f=0$)

In the reference model the rt location of the COG coincides with the rt location of the guiding pin, and the terms J_{rt} , J_{rf} and J_{tf} are equal to zero. Consequently, the focus displacement and the rotation about the focus axis are both neither dynamically (via the mass matrix) nor statically (via the stiffness matrix) coupled to either of the remaining DOF. The resulting focus and radial FRF are therefore perfectly straight lines without any resonance phenomena.

Influence of unbalance

Due to imperfections in the manufacturing process, slight deviations from the reference model are to be expected. To evaluate the influence of a slightly shifted location of the COG in the tangential direction, an extra mass of 50mg is added successively at the rt position of the counter mass and the lens (Fig.6.15). The location of the additional mass in the focus direction corresponds to that of the COG, and consequently the resulting principal axes of inertia still correspond to the r, t, and f directions.

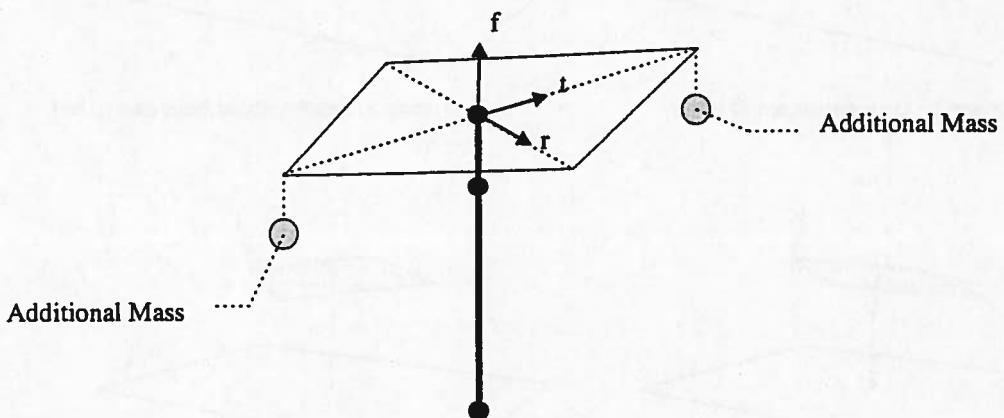


Figure 6.15 Location of extra mass

Due to the small extra mass and the resulting coupling between the various DOF of the actuator, the precise behaviour of the modes is affected, but the natural frequencies and the general shape are practically unchanged. However, in the focus and radial FRF of the two modified systems (Fig.6.16, Fig.6.17) significant changes

can be observed. In the balanced reference model, the node at which the focus force and radial torque is applied remains stationary in the four parasitic modes, and consequently (see Section 4.3.4) these modes were not visible in the FRF. However, in the unbalanced situation, that specific node is no longer stationary, and although its motion is relatively small compared to other nodes of the model, resonances can now be observed in the FRF :

- focus : 490 Hz (pitch mode)
- radial : 690 Hz (roll mode) and 3640 Hz (radial translation)

If the unbalance mass is located at the counter-mass side, all three resulting resonance phenomena lead to a zero/pole combination and a phase lead (Fig.6.16a+b). For the pitch mode (focus direction) and the roll mode (radial direction) this is a favourable situation, as both natural frequencies are in the medium-frequency range. The 3640Hz mode, which is visible in the radial FRF, could become a problem, as it is close to the transition-frequency from the medium- to the high-frequency range and one therefore has to rely on sufficient amplitude-margin. In the current simulation, which is based on 1% relative damping, the amplitude at the resonance is small enough but in view of the uncertainty regarding the real damping a larger margin is favourable.

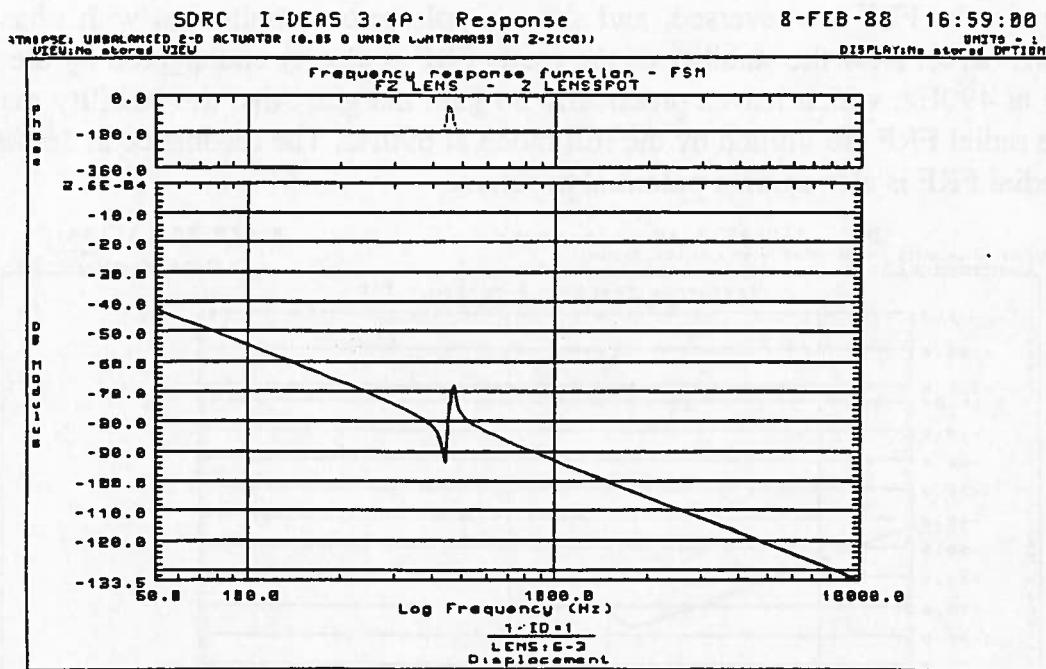


Figure 6.16a Focus FRF with extra mass at counter mass side ($r_t=r_{t,counter\ mass}$, $f=f_{coo}$)

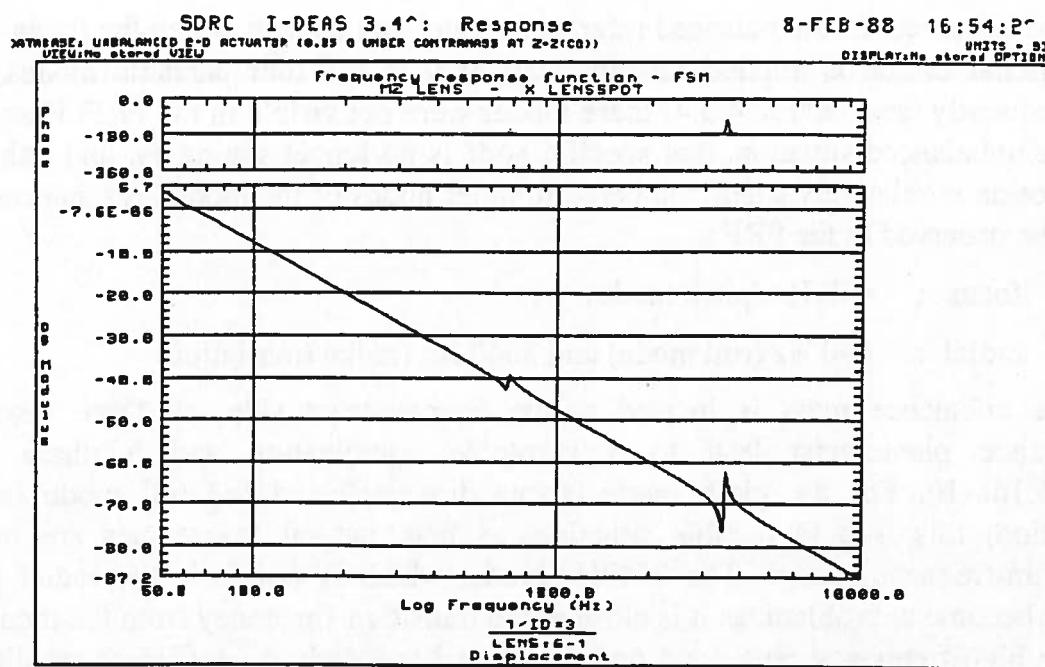


Figure 6.16b Radial FRF with extra mass at counter mass side ($r=r_{counter\ mass}$, $f=f_{COG}$)

When the unbalance mass is at the lens side, the characteristics of all three parasitic modes in the FRF are reversed, and show a pole/zero combination with phase lag (Fig.6.17a+b). Now the stability of the focus FRF is clearly endangered by the pitch mode at 490Hz, which leaves practically no gain margin. Also the stability margins of the radial FRF are limited by the roll mode at 690Hz. The resonance at 3640Hz in the radial FRF is also again a potential problem.

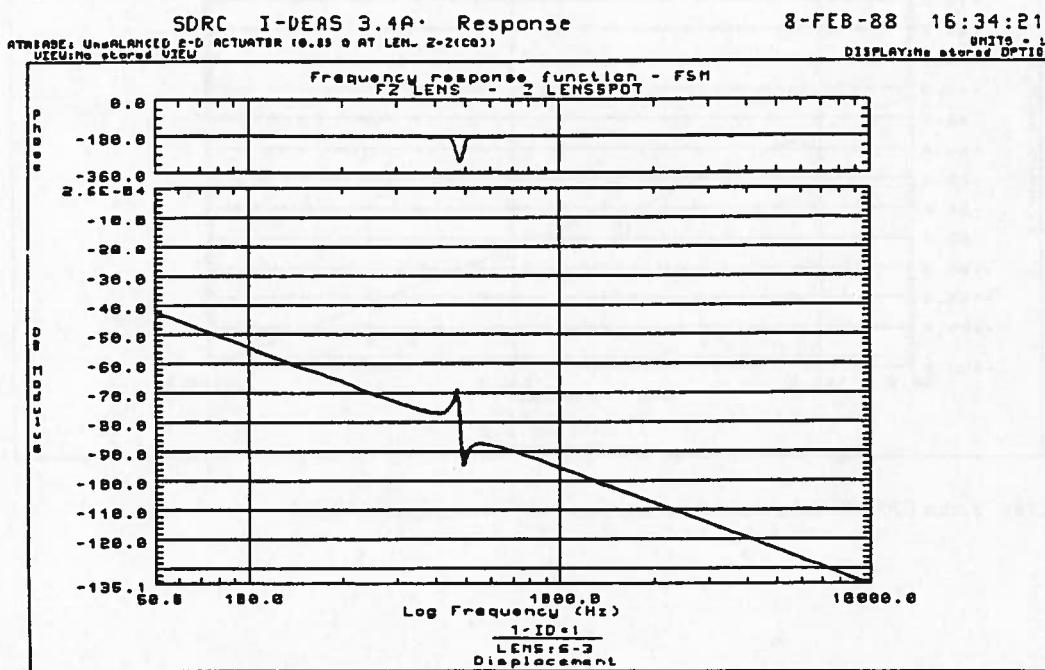


Figure 6.17a Focus FRF with extra mass at lens side ($r=r_{lens}$, $f=f_{COG}$)

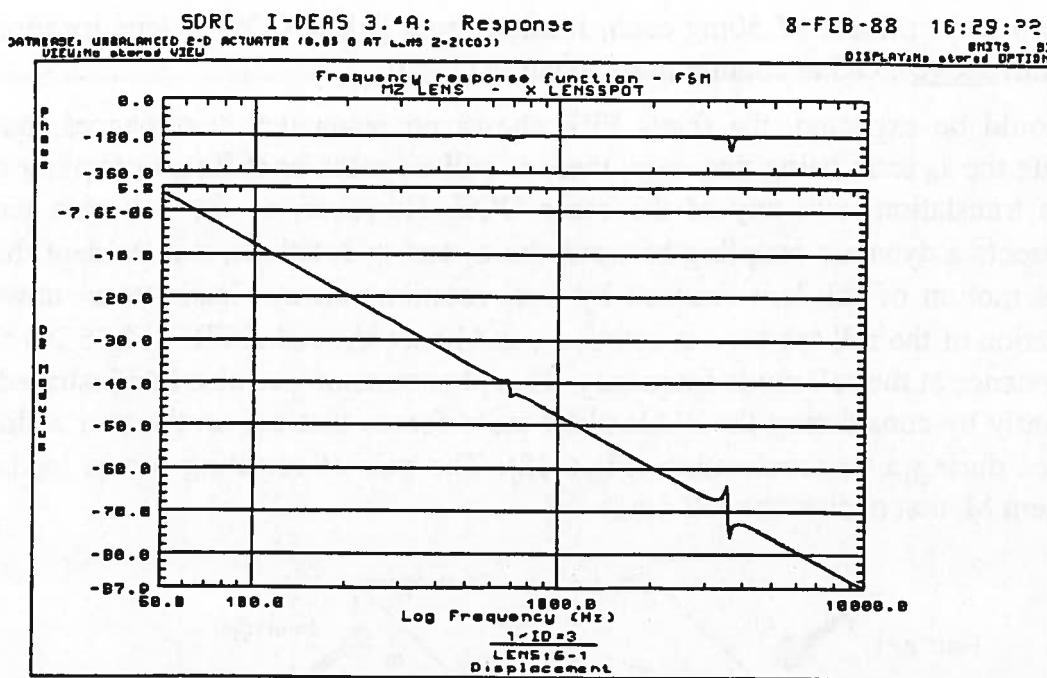


Figure 6.17b Radial FRF with extra mass at lens side ($r_t=r_{lens}$, $f=f_{COG}$)

On the basis of these simulations it becomes clear that accurate balancing of the actuator is very important, because even a small unbalance of 50mg on a total mass of the actuator of 1.25g can have a significant impact on the stability of the control loop. Furthermore, these calculations indicate that an unbalance mass located at the lens side results in resonances with a pole/zero (phase lag) behaviour in the FRF, whereas unbalance at the counter-mass side leads to a zero/pole (phase lead) characteristic.

Principle Moments do not coincide with r, t, and f directions

In the previous section it was concluded that the actuator needs to be properly balanced with respect to the location of the COG (static balancing about the guiding pin). Apart from tolerance aspects, this can be implemented relatively easily in the design by a proper location and inertia of the counter mass. From a design freedom point of view it is interesting to know whether the actuator should also be dynamically balanced (principal axes of inertia that coincide with the r, t, and f directions), because this would mean an extra constraint for the design.

To investigate the influence of the orientation of the principal axes of inertia, the following two situations are analysed, both of which lead to a term J_{tf} being non-zero :

1. Two extra masses of 50mg each, locate d 1mm above COG at lens location and 1mm below COG at counter-mass location ($J_{tf}>0$).

2. Two extra masses of 50mg each, located 1mm below COG at lens location and 1mm above COG at counter mass location ($J_{tf} < 0$).

As could be expected, the focus FRF shows no resonance disturbances, because despite the J_{tf} term being non-zero, there is still no mass or stiffness coupling of the focus translation with any of the other DOF. However, as the non-zero term J_{tf} represents a dynamic coupling between the φ_f and φ_t rotations, it is evident that the radial motion of the lens (caused by a φ_f rotation) always leads to an unwanted excitation of the roll mode (= φ_t rotation), and hence the radial FRF (Fig.6.19) shows a resonance at the roll mode frequency. This phenomenon can also be illustrated very elegantly by considering the d'Alembert mass forces that act on the two additional masses during a φ_f acceleration (Fig.6.18). The pair of resulting forces leads to a moment M_t that excites the roll mode.

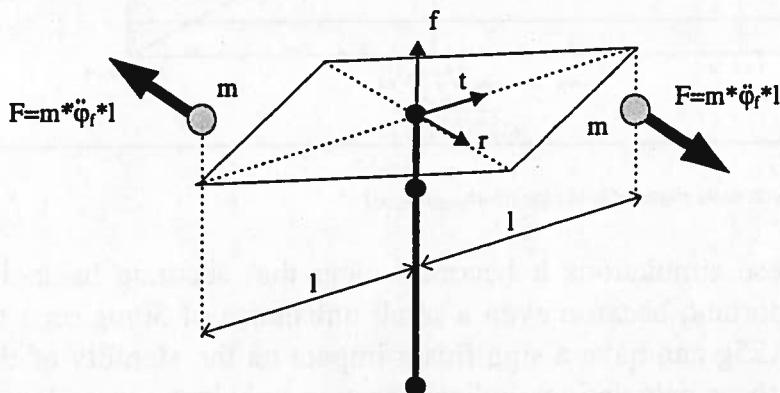


Figure 6.18 Inertia forces that result from a φ_f acceleration

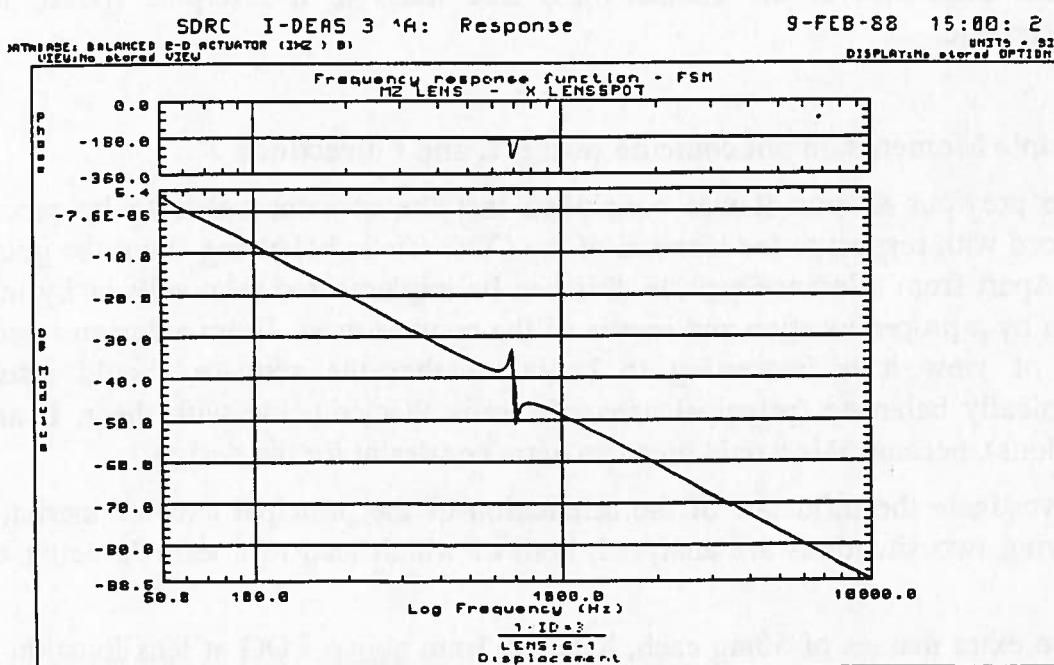


Figure 6.19a Radial FRF in case of $J_{tf} > 0$

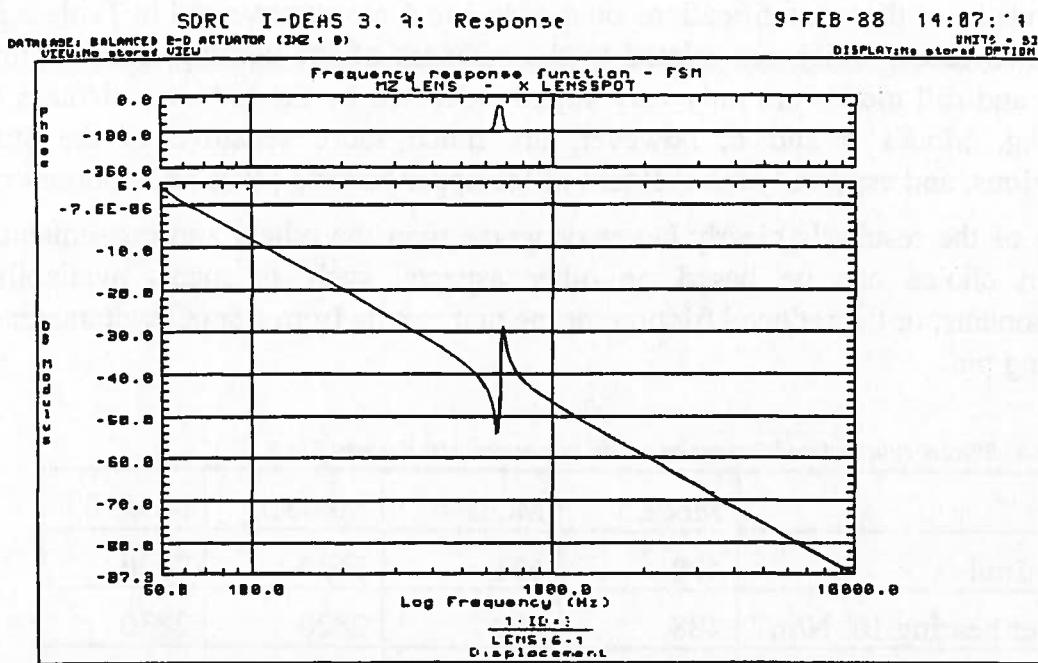


Figure 6.19b Radial FRF in case of $J_{tf}<0$

In the first case ($J_{tf}>0$) the radial FRF displays a pole/zero combination and a phase lag, whereas in the second case ($J_{tf}<0$) a zero/pole combination and a phase lead emerges. Referring again to the discussion of Chapter 4, and considering that the resonance frequency is located in the medium-frequency range, it is preferable to have the phase lead. Evidently, it is the ultimate goal not to excite the roll mode at all by a proper mass distribution, but if this is not possible due to design constraints, it is better to create a situation in with $J_{tf}<0$. Consequently it is better to place the counter mass higher than the lens than lower.

Influence of structural modifications on natural frequencies

The previous analyses have guided the design of the actuator body and its mass distribution from both a dynamics and stability point of view. Uncertainties still exist about the bearing stiffness, and various design choices related to the guiding pin still need to be made. Therefore a number of modifications are analysed :

- reduced stiffness upper bearing ($10^7 \text{ N/m} \Rightarrow 10^6 \text{ N/m}$)
- reduced stiffness lower bearing ($10^7 \text{ N/m} \Rightarrow 10^5 \text{ N/m}$)
- location upper bearing 2mm reduced in f direction
- reduced diameter guiding pin ($0.6\text{mm} \Rightarrow 0.5\text{mm}$)

The results of these modifications on modes 3 to 6 are summarised in Table 6.3. The first two cases, which are related to the stiffness of the bearing, indicate that the pitch and roll modes are only very slightly affected by the reduced stiffness of the bearing. Modes 5 and 6, however, are much more sensitive to the stiffness variations, and especially the stiffness of the upper bearing plays an important role.

None of the results is clearly better or worse than the others and consequently the design choice can be based on other aspects, such as space, availability of components, or the reduced friction torque that results from a smaller diameter of the guiding pin.

Table 6.3 Effect of structural modifications on natural frequencies [Hz] of modes 3 to 6

	Mode 3	Mode 4	Mode 5	Mode 6
Nominal	490	691	3630	3640
Upper bearing 10^6 N/m	488	688	2820	2830
Lower bearing 10^5 N/m	475	670	3580	3590
Upper bearing -2 mm	370	543	1590	1690
Diameter pin 0.5 mm	306	474	2510	2530

6.4.2. Overall Model

On the basis of the initial sketches of the compact disc module (Fig.6.9), a first 3D model [RAN88/2] was created long before the first prototypes were available. In this stage of the development, the size and weight of the various components, plus the layout of the drive, was only roughly known, and consequently the simulation model was also relatively rough and undetailed. The model (Fig.6.20), which was created in the System Dynamics module of the I-DEAS software (SDRC), consisted of :

- base plate : FE model (shell elements) imported from ANSYS
- disc motor : rigid body
- disc : 18 shell elements
- laser detection unit : 5 shell elements
- linear motor 1.stage : 1 shell element
- slide + coil : rigid body
- guiding of slide : 3 beam elements (each)
- 2D-actuator : rigid body
- guiding pin actuator : 2 beam elements

These components are linked either by rigid connections or by connectors which represent intermediate flexibilities and which enable the required motions of the slide and the 2D actuator.

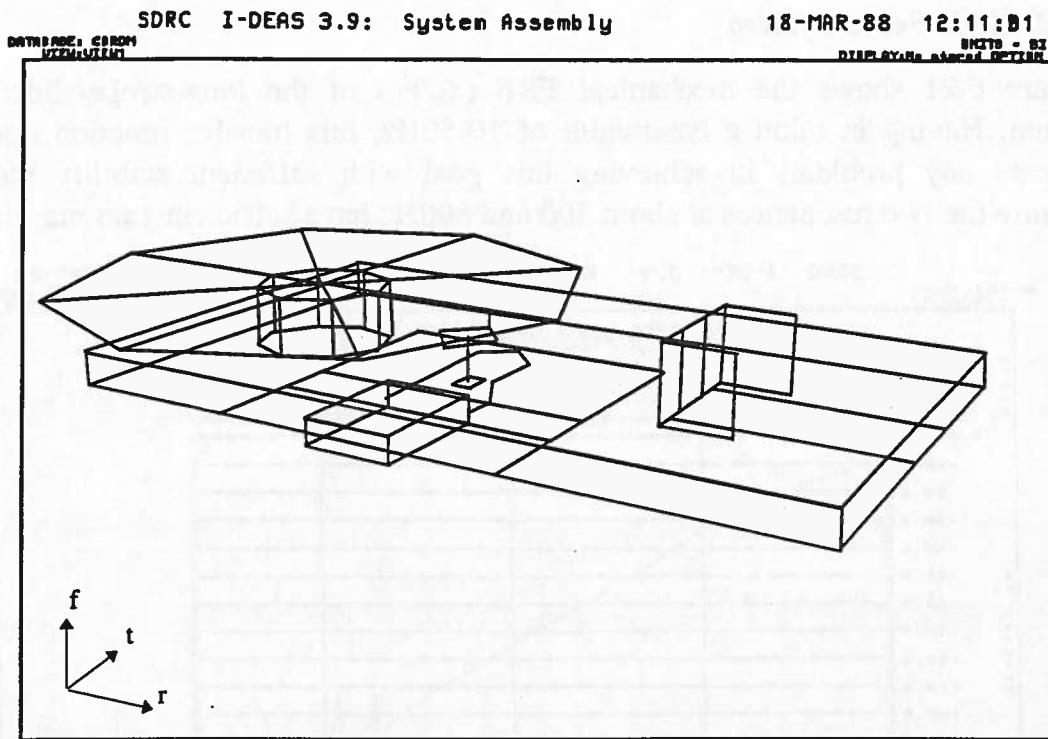


Figure 6.20 Model of the CDM-8.1

With the exception of the base plate and the disc, the various components were modelled as rigid bodies (no internal flexibility) or as FE components with very little detail. The base plate was modelled in more detail; first because its dynamics were considered to play a dominant role in any parasitic motion of the information track and the laser detection unit, and second because it required very little extra effort. The dynamics of the disc itself, which has its first modes at about 150Hz, are also crucial for the parasitic motion of the track and are therefore also included. As the disc is always part of the system, its properties are well-known and relatively easy to model.

To be able to judge the FRF in a realistic manner, user-defined DOF which incorporate the contribution of the motions of each component in the optical light path to the radial and focus servo signal have been introduced into the model. These 3D relations were derived according to the guide lines given in Appendix F.

The dynamic analysis of the entire system yields a large number of natural frequencies and modes (100 modes are required to describe the dynamic behaviour up to 7kHz). However, with respect to servo control and stability, only those modes that contribute significantly to the FRF of the radial or focus servo system are of

interest. After adding 1% relative damping to all modes, the mechanical FRF ($x_{\text{servo}}/F_{\text{servo}}$) was examined for each of the three servo loops.

6.4.3. Slide Servo System

Figure 6.21 shows the mechanical FRF (x_{ls}/F_{ls}) of the long-stroke-slide servo system. Having in mind a bandwidth of 10-50Hz, this transfer function does not indicate any problems in achieving this goal with sufficient stability margins, because the two resonances at about 500 and 800Hz have sufficient gain margin.

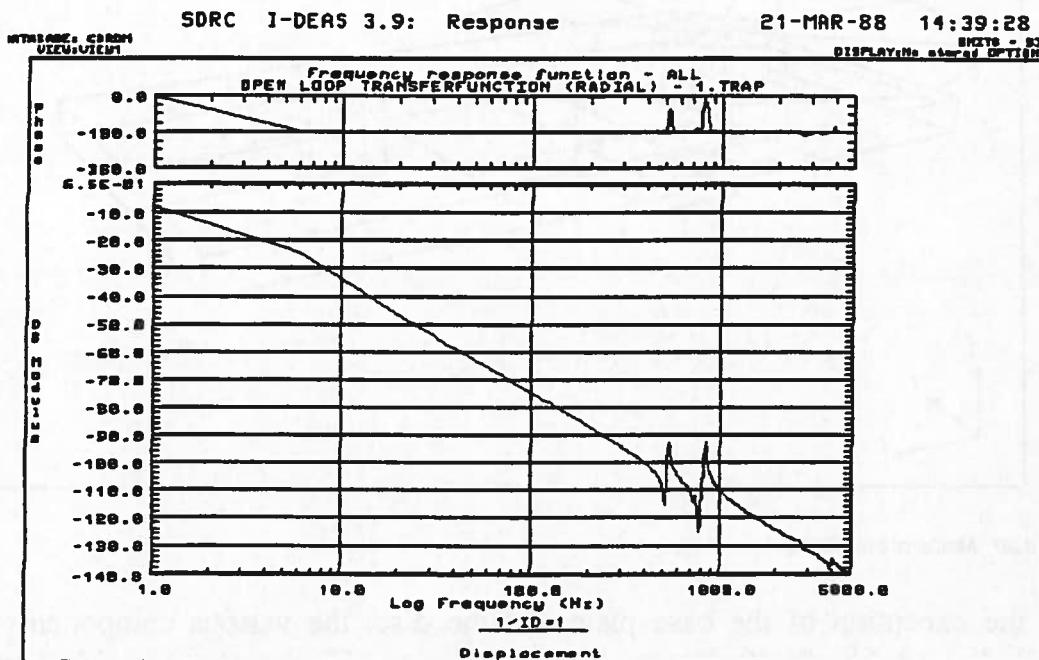


Figure 6.21 Mechanical FRF x_{ls}/F_{ls} of slide servo

Although there is no direct stability problem, it is still interesting to investigate the modes that correspond to the two peaks in the FRF. Without an animated display it is difficult to see, but the first mode (Fig.6.22) is dominated by a φ_t -rotation of the actuator, which shows good agreement with the roll-mode of an individual actuator, which was already found in the preliminary investigation (mode 4 in Fig.6.14). As this mode causes a small counter-motion of the slide in general, and more specifically of the coils at which the driving force is applied, it is unfortunately also visible in the slide control. The second mode (Fig.6.23) shows a φ_f -rotation of the entire slide due to the bending flexibility of the guidance, and initiates in turn a roll motion of the actuator.

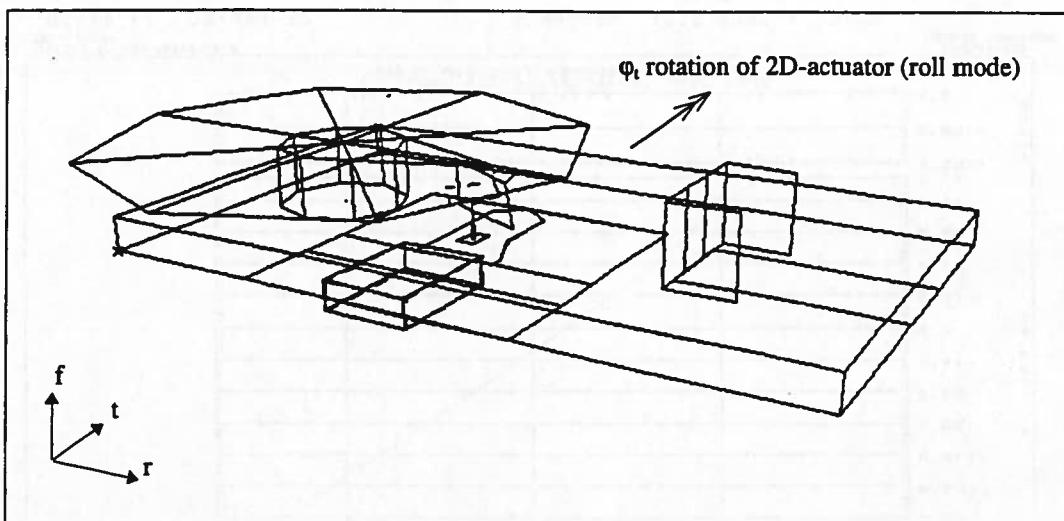


Figure 6.22 Mode 22 (529 Hz)

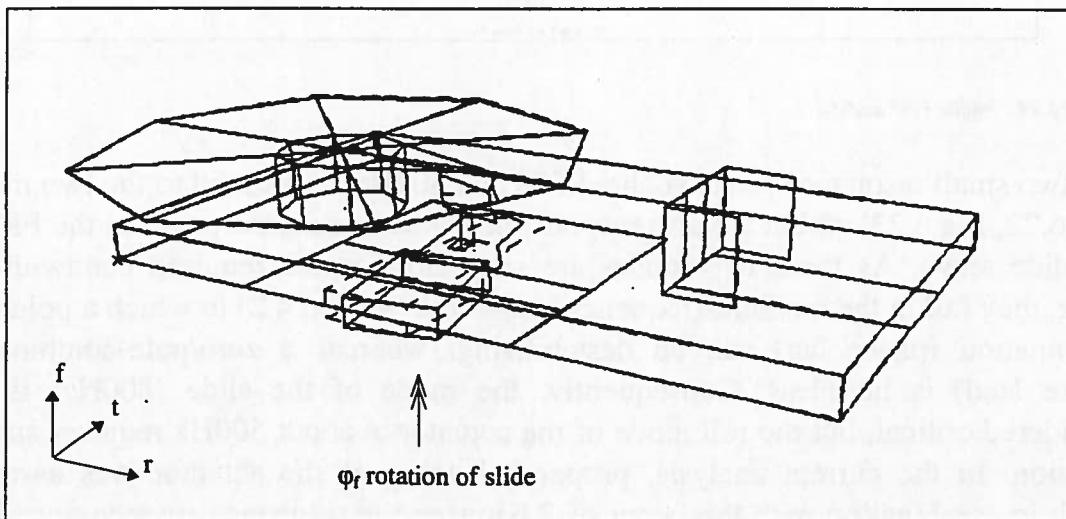


Figure 6.23 Mode 29 (837 Hz)

The main reason for the occurrence of the second mode in the FRF is the fact that the slide is not driven in its centre of gravity. Because this mode and the first mode are closely related (both display a φ_f rotation of the slide) it is plausible that driving the slide in its centre of gravity will also have a positive effect on the excitation of the first mode.

6.4.4. Radial Servo System

Figure 6.24 shows the radial mechanical FRF (x_{ss}/M_{ss}) of the short-stroke servo, in which x_{ss} represents the radial servo position (including the complete optical relation), and M_{ss} stands for the driving torque that acts between the actuator and the slide.

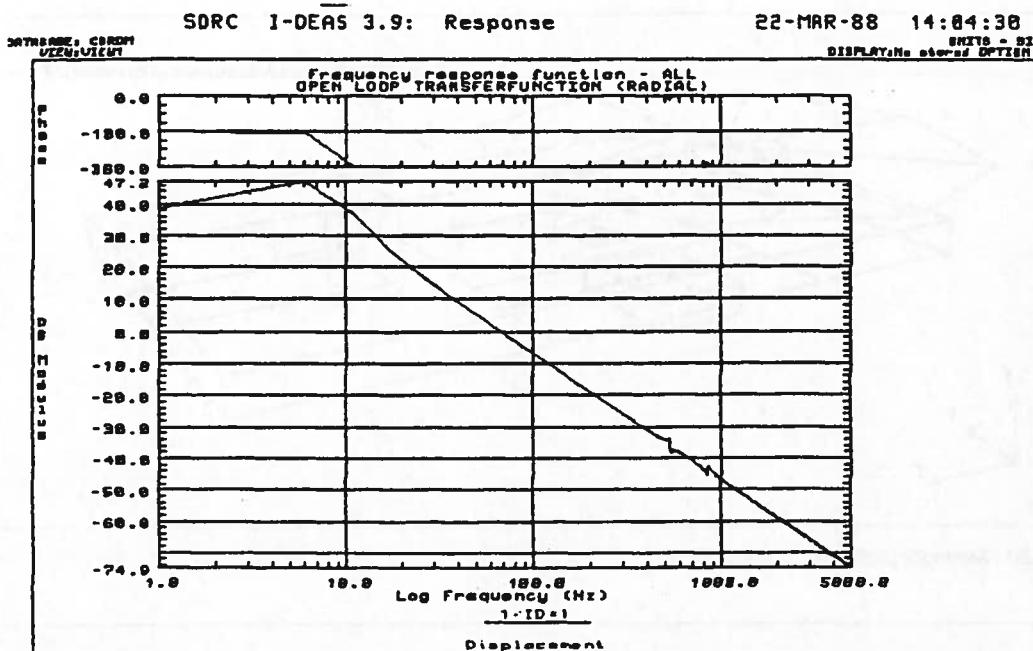


Figure 6.24 Radial FRF X_{ss}/M_{ss}

The two small resonance peaks at about 500 and 800Hz correspond to the two modes (Fig.6.22, Fig.6.23) which are also responsible for the resonance peaks in the FRF of the slide servo. As these two modes are very close to the required bandwidth of 1kHz, they fall in the medium-frequency range (see Section 4.2) in which a pole/zero combination (phase lag) can be destabilising, whereas a zero/pole-combination (phase lead) is harmless. Consequently, the mode of the slide (800Hz) is not considered critical, but the roll mode of the actuator at about 500Hz requires special attention. In the current analysis, proper balancing of the actuator was assumed which in combination with the factor of 2 between the resonance frequency and the bandwidth frequency results in a limited destabilisation of the servo control-loop. However, it is again shown that proper balancing of the actuator is an important issue. Furthermore it is also important that the frequency of the roll mode should not be increased, because the current phase-shift of -30° can be tolerated at about 500Hz (because of sufficient gain margin) but at the bandwidth frequency of 1kHz it is absolutely intolerable (no gain margin exists at the bandwidth frequency).

As the slide plays such an important role in the dynamics so far, a refinement of the slide model seems essential as soon as more detailed information is available.

In search for possible remedies to increase the robustness of the design a more conceptual model is derived to understand the mechanism behind the appearance of these two modes in the FRF. Figure 6.25 shows a planar model of the slide with its guiding system and with the actuator on top of it. The actuator is free to rotate with respect to the slide, whereas the relative displacements are restrained by translational springs. The X mode of the actuator with respect to the slide in this planar model can

be considered to have a similar effect as the ϕ_t mode (roll mode) of the actuator in the 3D model (Fig.6.22).

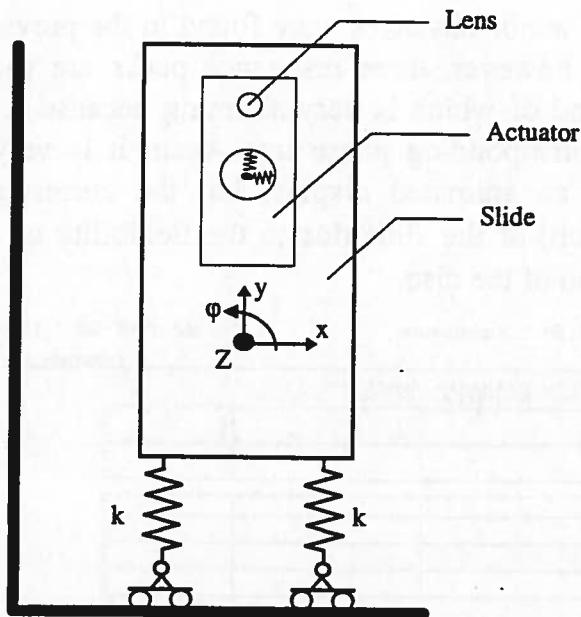


Figure 6.25 Conceptual model used to understand the mechanisms behind the resonance peaks in the radial FRF

The system consists of two components each with three DOF, so there are six modes in total. Assuming that the system is symmetric about the Y axis, the two modes in Y direction are completely decoupled from the other modes, and need not be considered. Furthermore, there are two 0 Hz modes that correspond to the desired motion, namely movement of the slide in X direction, and rotation ϕ of the actuator around its guiding pin. The remaining two modes are the ones of interest. The first is dominated by a vibration of the actuator in X direction, which evidently also leads to a certain amount of rotation and displacement of the slide. The second consists mainly of a rotation of the slide, which is only little influenced by the small mass of the actuator. Because this rotation also leads to an X displacement of the guiding pin, the dynamics of the actuator are also addressed.

In seeking a possible solution to the problem of the two resonance peaks in the radial loop of the second stage, it is essential to realise that the rotation of the slide itself is not the problem. The main problem is the fact that this rotation also leads to a displacement of the guiding pin of the actuator. On the basis of this conception the solution is quite evident: the position of the guiding pin of the actuator should coincide with the COG of the slide, which is the node of the slide-rotation mode. If this is the case, the two modes become decoupled with respect to the contribution of $x_{actuator}$ and ϕ_{slide} . This implies that the reaction torque applied to the slide, when driving the desired rotation of the actuator, still excites the rotational mode of the slide, but this does not lead to a displacement of the guiding pin, and consequently the mode of the actuator is not excited.

6.4.5. Focus Servo System

Despite the concern about the influence of the roll mode of the actuator on the stability of the radial loop, no major obstacles were found in the previous analyses. In the focus FRF (Fig.6.26), however, three resonance peaks are found near the bandwidth of 1kHz, the second of which is very alarming because it introduces a pole/zero combination and corresponding phase lag. Again it is very difficult to interpret the results without an animated display, but the corresponding mode (Fig.6.27) is a φ_r rotation (pitch) of the slide due to the flexibility of the guidance combined with a φ_t deformation of the disc.

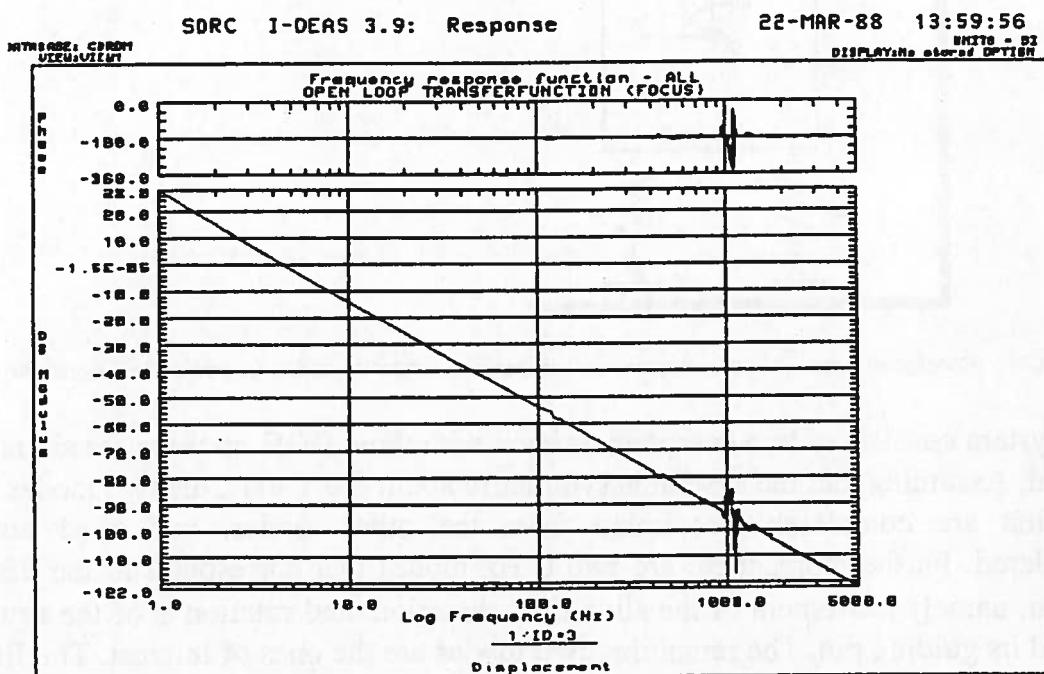


Figure 6.26 Focus FRF

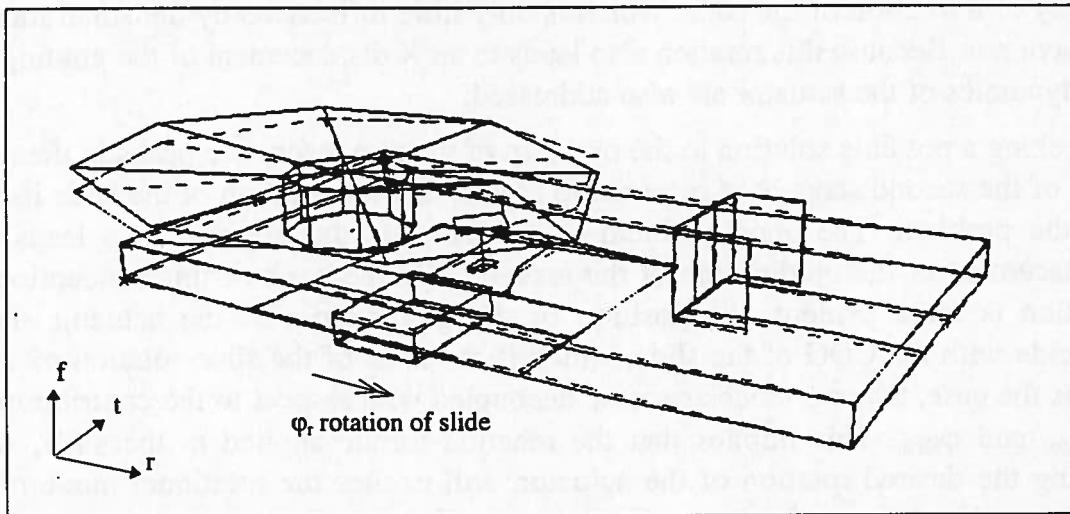


Figure 6.27 Mode at 1130 Hz

The severity of this mode in the FRF is somewhat unexpected because of the large mass ratio between the base plate (200gr) and the actuator (1.25gr). Apparently the underlying phenomenon cannot be explained on the basis of the simple conceptual model (Fig.6.4) used in the concept evaluation phase. The most distinct difference between the conceptual model and the observed phenomenon lies in the limited stiffness with which the slide is fixed to the base plate. Extending the model with an extra mass and stiffness that represents the slide and its connection to the base plate leads to the model shown in Fig.6.28.

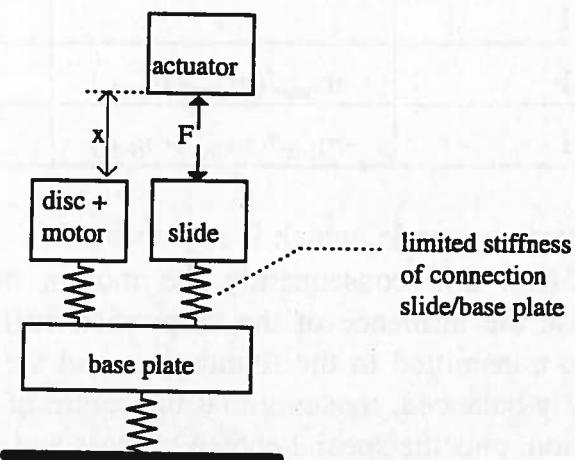


Figure 6.28 Extended concept model of focus servo system including limited stiffness of slide and guidance

With this extended concept model the observed phenomenon can readily be demonstrated (Fig.6.29) by choosing the mass and stiffness properties of the slide and the disc (+motor) such that for a fixed base plate the two resulting natural frequencies of disc and slide coincide.

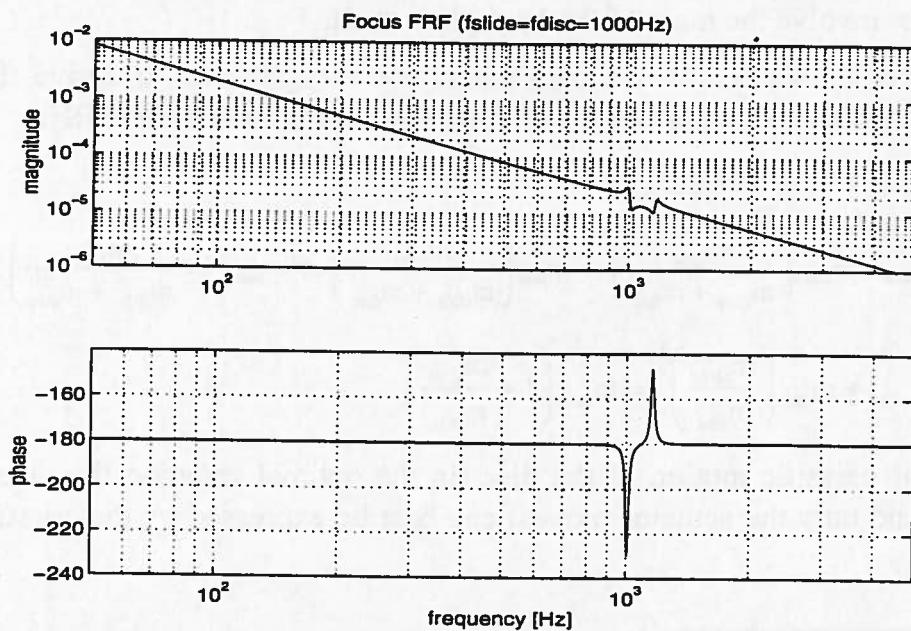


Figure 6.29 Focus FRF (x/F) corresponding to the extended concept model of Fig.6.28 ($f_{slide} = f_{disc} = 1000\text{Hz}$)

Due to the special choice of mass and stiffness properties, the mode-shapes corresponding to the extended model satisfy the conditions in Table 6.4 (influence of suspension stiffness is negligible).

Table 6.4 Mode-shapes corresponding to the extended concept model of Fig.6.28.

	Mode 1 (suspension mode)	Mode 2 (1. internal mode)	Mode 3 (2. internal mode)
$\Phi_{i,\text{base}}$	1	1	0
$\Phi_{i,\text{slide}}$	1	$-m_{\text{base}}/(m_{\text{disc}}+m_{\text{slide}})$	1
$\Phi_{i,\text{disc}}$	1	$-m_{\text{base}}/(m_{\text{disc}}+m_{\text{slide}})$	$-m_{\text{slide}}/m_{\text{disc}}$

The first mode is a suspension mode, which is not visible in the FRF due to its low frequency ($f_{\text{suspension}}=25\text{Hz}$), and consequently the motion of all three DOF is practically “1”⁸. Because the influence of the suspension stiffness is negligible at 1kHz, no forces can be transmitted to the foundation, and the motion of the three masses must be internally balanced, meaning that the centre of gravity must remain unchanged. This condition, plus the special choice of mass and stiffness of slide and disc, leaves only two possible modes (with theoretically identical frequency) :

Mode 2 (1. internal mode) is characterised by the fact that the slide and the disc have the same amplitude and phase, whereas the base plate moves in the opposite direction. The amplitude ratio is such that the centre of gravity does not change.

At mode 3 (2. internal mode), the base plate remains stationary, and disc and slide move in opposite directions. Their amplitude ratio is determined by m_{disc} and m_{slide} and does not involve the mass of the base plate at all.

Based on the knowledge of the modes it is not very difficult to derive the modal mass of each of these modes and to calculate the contribution to the FRF .

$$m_1 = m_{\text{base}} + m_{\text{slide}} + m_{\text{disc}} \quad (6.11a)$$

$$m_2 = m_{\text{base}} + m_{\text{slide}} \left(\frac{m_{\text{base}}}{m_{\text{slide}} + m_{\text{disc}}} \right)^2 + m_{\text{disc}} \left(\frac{m_{\text{base}}}{m_{\text{slide}} + m_{\text{disc}}} \right)^2 = m_{\text{base}} \left(1 + \frac{m_{\text{base}}}{m_{\text{slide}} + m_{\text{disc}}} \right) \quad (6.11b)$$

$$m_3 = m_{\text{slide}} + m_{\text{disc}} \left(\frac{m_{\text{slide}}}{m_{\text{disc}}} \right)^2 = m_{\text{slide}} \left(1 + \frac{m_{\text{slide}}}{m_{\text{disc}}} \right) \quad (6.11c)$$

The level of parasitic motion of the disc (in the optimal situation the disc remains stationary and only the actuator moves) can best be expressed by the co-efficient of

⁸ This is not exactly true: however, small internal deformations are of no interest in this analysis.

the high frequency modal contribution, which can directly be related to the mass of the actuator :

$$\frac{\phi_{1,slide} \phi_{1,disc}}{m_1} = \frac{1}{m_{base} + m_{slide} + m_{disc}} \quad (6.12a)$$

$$\frac{\phi_{2,slide} \phi_{2,disc}}{m_2} = \frac{\left(\frac{m_{base}}{m_{slide} + m_{disc}} \right)^2}{m_{base} \left(1 + \frac{m_{base}}{m_{slide} + m_{disc}} \right)} = \frac{1}{\left(m_{slide} + m_{disc} \right) \left(1 + \frac{m_{slide} + m_{disc}}{m_{base}} \right)} \quad (6.12b)$$

$$\frac{\phi_{3,slide} \phi_{3,disc}}{m_3} = \frac{1 \left(-\frac{m_{slide}}{m_{disc}} \right)}{m_{slide} \left(1 + \frac{m_{slide}}{m_{disc}} \right)} = -\frac{1}{m_{slide} + m_{disc}} \quad (6.12c)$$

A closer look at these contributions reveals that the influence of the mass of the base plate is either absent (mode 3) or has an effect which is the opposite from the expected (contribution of mode 2 increases with increasing values of m_{base}). As a result of this, the design no longer profits from the huge mass ratio between the actuator and the base.

The severity of these modes is drastically reduced as soon as the natural frequencies of the slide and the disc no longer coincide. Figure 6.30 shows the FRF after reducing the natural frequency of the slide to 500Hz⁹.

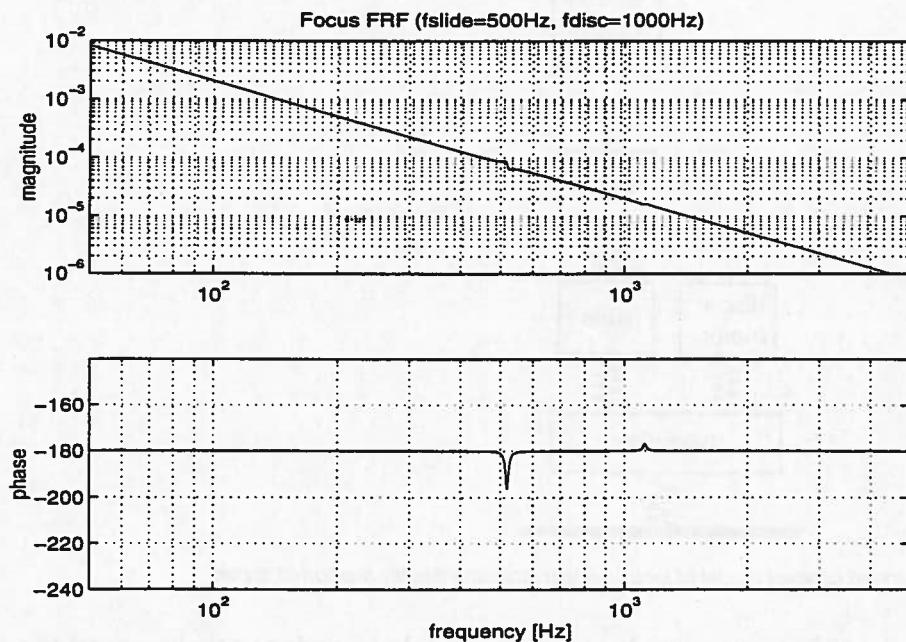


Figure 6.30 Focus FRF (x/F) corresponding to the extended concept model of Fig.6.28 ($f_{slide}=500\text{Hz}$, $f_{disc}=1000\text{Hz}$)

⁹ Natural frequency of the slide under the assumption of a fixed base plate.

This approach seems interesting, but one can only be certain to avoid natural frequencies involving the compact disc if the slide frequency is chosen significantly lower than the first natural frequency of a clamped compact disc (about 150Hz). However, choosing such a low natural frequency and corresponding low stiffness of the slide support is unacceptable from an accuracy point of view, because the resulting static deflections of the slide due to gravity effects would become too large (a natural frequency of 100Hz already results in a static deflection of $25\mu\text{m}$). This would endanger the correct orientation of the guiding pin of the 2D actuator relative to the disc motor, and also the orientation of the corner mirror relative to the laser detection unit.

Alternatively, one could also create a mechanical filter by introducing flexibility in the connection between the slide and the stator of the Lorenz motor that drives the actuator (Fig.6.31). This approach leads to a significant reduction of the high-frequency content of the reaction forces that are transmitted via the stator to the slide, and also a reduction of force that is transmitted via the base to the disc.

In contrast to the slide, the stator does not hold any delicate components, so a natural frequency of about 75Hz is acceptable. This frequency is sufficiently low (compared with the lowest natural frequency of the disc) and it is also high enough not to interfere with the suspension frequency of the complete unit or the rotational speed of the disc. The effect of such a mechanical filter with a natural frequency of 75Hz on the FRF is shown in Fig.6.32. As one can see, neither the mode of the slide nor the mode of the disc (+motor) is visible in the FRF. There is a new resonance at 75Hz, but this does not introduce instability.

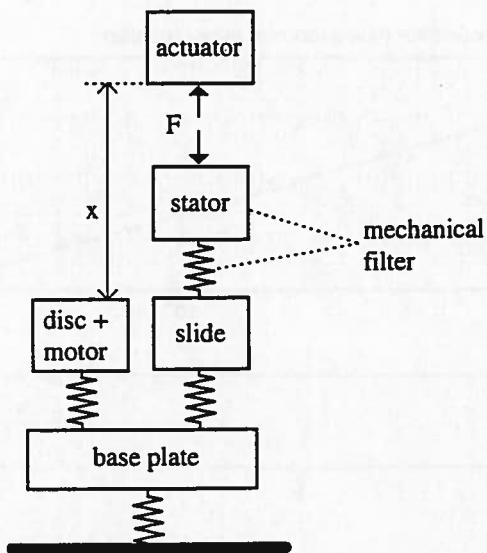


Figure 6.31 Extended concept model of focus system including flexibly supported stator

After this excursion into very elementary models in order to understand the observed phenomenon and to show the feasibility of a mechanical filter, the concept was also included in the 3D system model; it performed according to expectation.

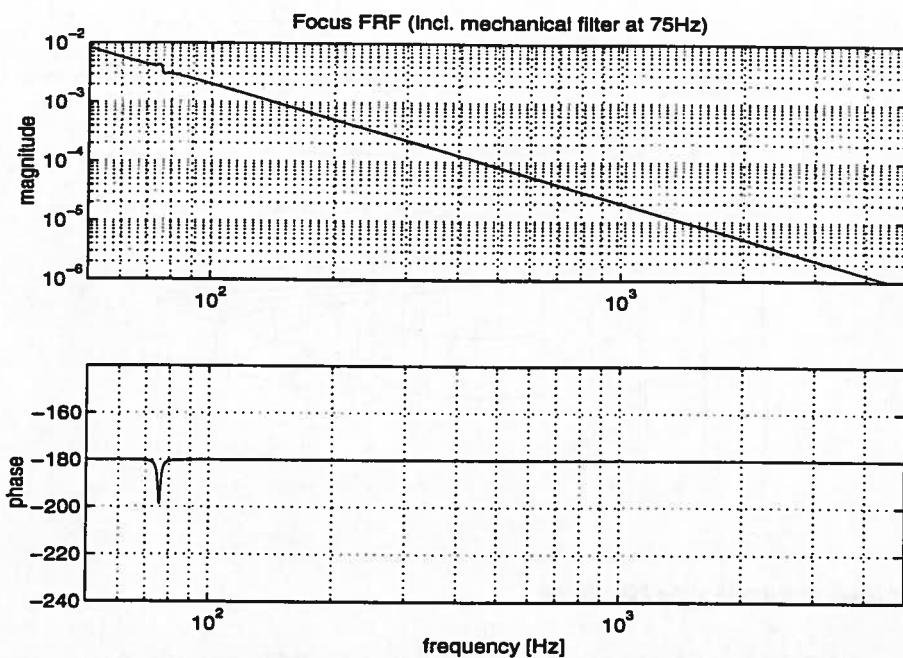


Figure 6.32 FRF corresponding to the extended concept model of Fig.6.31 ($f_{\text{stator}}=75\text{Hz}$)

6.5. Component Evaluation

In previous analyses, the actuator body has been considered rigid, which is only true in a limited frequency range. In fact the actuator body is probably the most critical component of the entire system, because any dynamics between the location of the driving force and the lens is visible in the frequency response as a -2 slope/pole/-4 slope characteristic. In Section 4.3.1 it was explained that this type of resonance behaviour can only be tolerated from a stability point of view if the natural frequency is sufficiently high. Aiming at a servo bandwidth f_b of 1 kHz leads to the following target for the lowest internal natural frequency f_e of the actuator (see design rules for actuator flexibility in Section 4.3.1) :

$$f_e > (5 \cdot 10) f_b = 5 \cdot 10 \text{ kHz} \Rightarrow \text{target} = 10 \text{ kHz}$$

The previous analysis of the entire system revealed that the contribution of the dynamics of the entire module to the FRF is mainly found in the 1kHz region. In the 10kHz frequency range the contribution of any undesired dynamics of the module to the servo position signal is negligible. The signal is completely dominated by the dynamics of the actuator. As a consequence of this, the actuator can be analysed separately without including the effect of the other components. Furthermore, at 10kHz the stiffness with which the actuator is guided is negligible compared to the mass effects, and therefore it is analysed as a free-free system. The finite element model [SCH88/1] of the initial actuator design is shown in Fig.6.33.

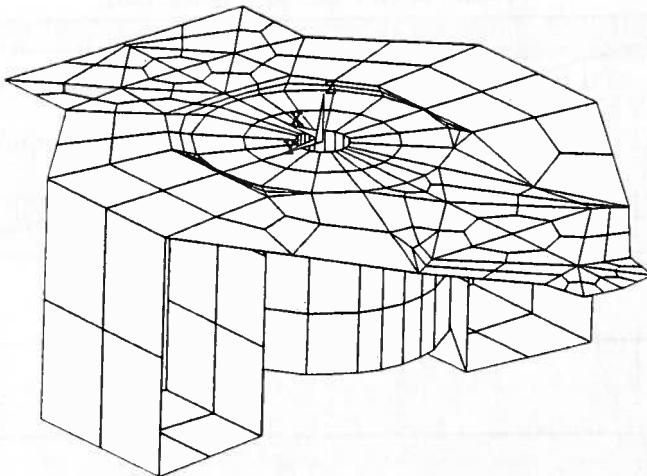


Figure 6.33 Finite element model of the 2D actuator

The entire model is based on shell elements with different thickness and material properties. Equivalent values for the thickness and material properties were used to approximate the coil properties, and to include the effect of ribs. The effect of glue between the focus coil and the radial coil was approximated by connecting nodes of the focus coil with nodes of the radial coil via stiff beam elements.

The lowest internal mode was found at 7.3kHz, whereas the second and third natural frequencies amounted to 11.2kHz and 16.1kHz. The dynamics of these three modes is visible in the focus and the radial FRF, but in the focus direction the contribution is more dominant. Figure 6.34 shows the mechanical focus FRF which is generated by applying focus forces at the outline of the focus coil, and calculating the focus displacement of the lens. As expected, one can observe a -2 slope at low frequencies, and a -4 slope above the resonances¹⁰.

Despite the fact that the mechanical design target of 10kHz was not met, a bandwidth of 1kHz could still be realised (see Fig.6.35) in combination with the following elements in the control loop :

- integrator up to 100Hz
- lead filter between 450Hz and 4800Hz
- voltage amplifier (the actuator is driven by a voltage amplifier, which implies that the electrical time constant of the actuator ($=20\mu s$) adds an extra first order low-pass filter to the control loop at about 8kHz).

¹⁰ It is interesting to note that it is actually the combined contribution of the three modes that results in the high frequency decoupling, or -4 slope, of the lens. The first mode by itself has an α value (see equations (4.3) and (4.4) for the definition of α) that equals -0.675, and would therefore lead to a FRF of the type “-2slope/pole/zero/-2slope”. However, the summed contribution of the three modes ($\alpha_1=-0.675$, $\alpha_2=-0.4$, $\alpha_3=0.075$) equals -1, and explains the decoupling above the frequency of the third mode.

The main reason for the stability of the open loop is the fact that the two dominant resonances have a favourable phase behaviour. In combination with the phase lag introduced by the controller and the actuator time constant, the resonances lead to amplitude excursions in a safe region of the Nyquist diagram.

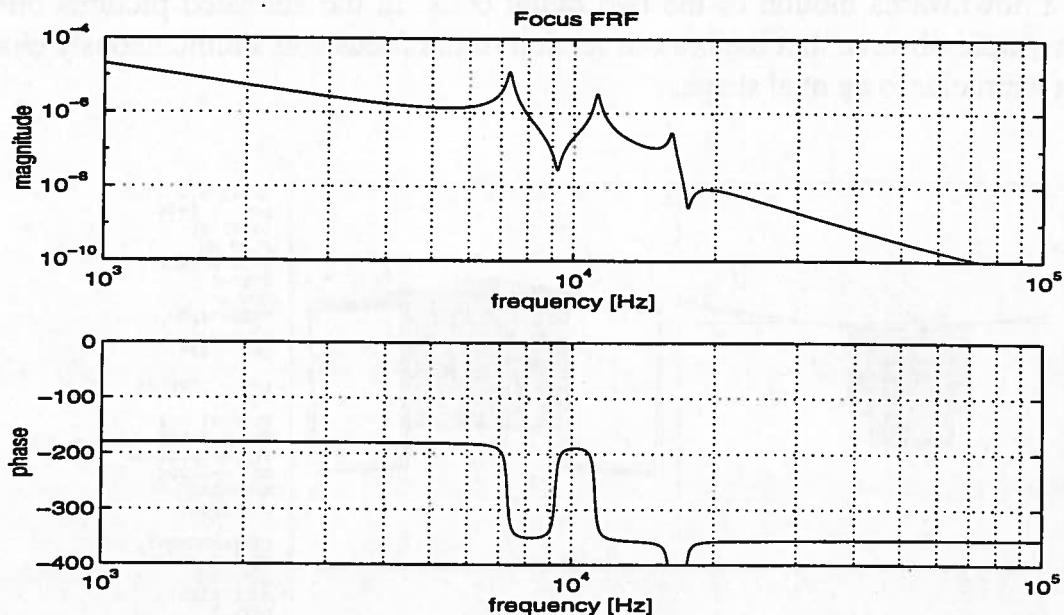


Figure 6.34 Mechanical focus FRF ($u_{\text{lens}}/F_{\text{zcoil}}$) of actuator

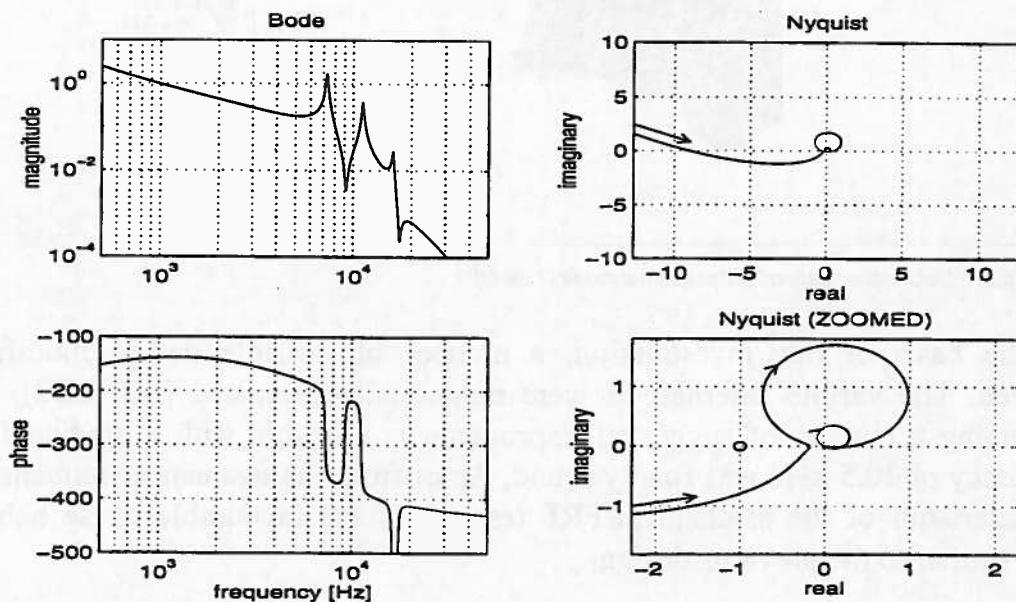


Figure 6.35 Open-loop focus FRF

Although the required bandwidth of 1kHz was achieved with the current actuator design, an attempt was made to realise the initial target of 10kHz, because this would lead to a more robust design in terms of servo stability, and would offer more buffer with respect to tighter specifications in future designs.

As a starting point for the generation of alternatives, the mode-shape of the lowest internal mode (Fig.6.36) was investigated intensively so as to understand the deformation pattern and to find regions with large deformations. The mode is best characterised as a saddle-shape with an upwards motion of lens and counter mass, and a downwards motion of the two radial coils. In the animated pictures one can furthermore observe that the bottom section of the focus coil simultaneously changes from a circular to an oval shape.

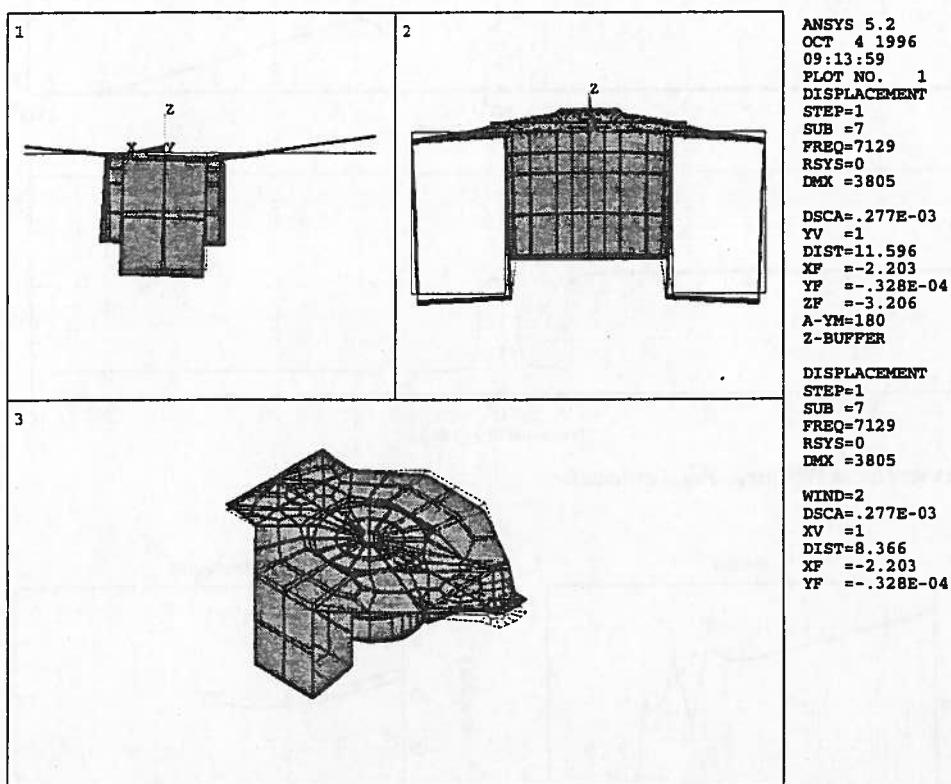


Figure 6.36 Deformation pattern of first internal mode (7.3kHz)

On the basis of this investigation, a number of possible design modifications evolved. The various alternatives were numerically evaluated [SCH88/2], and by combining a number of suggested improvements a design with a predicted lowest frequency of 10.5 kHz was finally found. Apart from the resonance frequencies, the characteristics of the mechanical FRF (especially the favourable phase behaviour) were similar to the previous design.

Tests of the prototype revealed a resonance at about 9.5kHz, which is considered to be in good agreement with the model. A difference of less than 10-20% between the calculated and measured natural frequencies is not to be expected because one always has to deal with uncertainties in the model. Furthermore, a perfect match between the model and reality is not a target in itself. The main goal it to support the design process and to guide design decisions such that a better product is created.

7. Conclusions

Machine dynamics, and the interaction with the control system, plays a dominant role in the performance of fast and accurate servo-controlled positioning devices such as compact disc drives, wafer-steppers, and component-mounters.

“Modal analysis” is a numerical and experimental tool that can be very profitable in understanding the nature of complicated mechanical resonances. The mathematics of a single decoupled “modal” equation of motion can be translated into a graphical representation including all relevant data, which simplifies the understanding and creative use of the modal concept. The introduction of the terms “effective” modal mass and stiffness enables a unique link between the modal and the physical domain.

From a servo stability point of view it is essential to investigate the mechanical FRF ($x_{\text{servo}}/F_{\text{servo}}$) which characterises the dynamic properties of the mechanical system. Once the dynamics of one individual mode is fully understood it is straightforward to construct this FRF and the interaction between the desired rigid body motion and the contribution of one additional mode. A closer investigation of this interaction reveals that only four¹ interaction patterns exist. The destabilising effect of a mechanical resonance depends not only on the resulting typical interaction pattern in the FRF, but also on its frequency in relation to the intended bandwidth frequency of the control loop. On the basis of these stability considerations, design guidelines for the mechanics of a servo positioning devices are derived, so as to minimise the effect of mechanical vibrations on the stability of the controlled system.

In view of its importance to the overall performance, the effect of machine dynamics should be monitored during the entire design process through the use of modelling and simulation techniques. However, it is vital for the success of modelling and simulation as a tool to support design decisions, that analysis data are translated into useful information, and that this information is available on time. This requires a proper balance between accuracy and speed that can best be achieved by a top-down analysis process, which is closely linked to the phases in the design process, and in which the simulation models are step-wise refined. As an example, the step-wise modelling approach used during the development of the CDM-8 compact disc module is presented in detail.

¹ One of these four interaction patterns has been excluded from further analysis because it corresponds to a non-minimum phase system, which should not occur in a well-designed device.

When many parts of the mechanical system need to be modelled in great detail it is not advisable to build one, single, huge FE model (error-prone and time/hardware consuming) but rather to apply so-called “sub-structuring” techniques. The Craig-Bampton approach, which is a component mode technique based on a combination of all boundary constraint modes plus a limited number of fixed interface normal modes, was found to be favourable. It has static solution capacity, and the frequency of the highest fixed-interface normal mode gives a good indication of the frequency range up to which the overall system results are valid.

Through the enormous performance drive in mechatronics systems, much has been learned in the past years about the influence of machine dynamics in servo positioning-devices. The lack of complicated mathematics in this thesis shows that these important aspects can be described and understood using operational knowledge of basic machine dynamics and control issues only. Therefore, the contents of this thesis is considered easily transferable knowledge, which hopefully will be included in the education of mechatronics engineers in the near future.

References

- BEU96 Beurden, M. van, "Frame vibrations and settling-time of servo systems", Master Thesis, Philips-CFT report CTB595-96-3047
- BOU83 Bouwens, H.J.J., "Electromechanical Position Servos", Philips-CFT, internal report, 1983
- COO79 Cool, J.C., Schijff, F.J. and Viersma, T.J. , "Regeltechniek", Elsevier, 1979
- CRA68 Craig, Jr., R.R. and Bampton, M.C.C., "Coupling of Substructures for Dynamic Analysis", AIAA Journal, 1968, Vol.6, No.7, pp.1313-1319
- CRA81 Craig, Jr., R.R., "Structural Dynamics - An Introduction to Computer Methods", John Wiley & Sons Inc., New York, 1981
- CRA87 Craig, Jr., R.R., "A review of time-domain and frequency-domain component mode synthesis methods", Int J. of Analytical and Experimental Modal Analysis, April 1987, pp.59-72
- DOF86 Dorf, R.C., "Modern control systems", Addison-Wesley Publishing Company, 1986
- DOR86 Doren, E. van, "Radiële en verticale volgsystemen voor een nieuwe generatie compact disc spelers", Philips-Consumer Electronics, internal report, 1986
- EIJ88 Eijk, J. van and Rankers, A.M., "Modal Analysis - The Philips approach towards an acceptable tool", 13th Int. Seminar on Modal Analysis. Leuven, 1988
- EIJ89/1 Eijk, J. van and Rankers, A.M., "Predictive Modelling - from research to engineering application", Int. conf. on advanced mechatronics, 1989, Tokyo
- EIJ89/2 Eijk, J. van (editor), "Predictive Modelling in Compact Disc Drives (Part I-IV)", Philips-Consumer Electronics, Optical Recording Lab., internal reports AR-46-236 (237,238 and 239), 1989
- EWI84 Ewins, D.J., "Modal Testing : Theory and Practice", Research Studies Press Ltd, England, 1984

- FRA88 Franklin, G.F., Powell, J.D. and Emami-Naeini, A., "Feedback control of dynamic systems", Addison-Wesley Publishing Company, 1988
- GIE92/1 Gielens, C., "Controlled mechanical systems in SDRC/I-DEAS System Dynamic Analysis", master thesis University Eindhoven, Philips CFT report nr. CTB 545.92.3022, 1992
- GIE92/2 Gielens, C., "Reductie aantal fysische vrijheidsgraden van modale componenten", University Eindhoven, Department Mechanical Engineering, Section CAE-systems, 1992
- GRO91 Groenhuis, H., "A design tool for electromechanical servo systems", PhD-thesis University Twente, WIBRO dissertatiedrukkerij Helmond, 1991
- HAR56 Hartog, J.P. den, "Mechanical Vibrations", McGraw-Hill, New York, 1956
- HEE72 Heesewijk, A.P.C. van, "Het werken met modellen", De Constructeur, Nr. 12, 1972
- HEI94 Heijden, H.L.M. van der, "Procedure voor de transformatie van dynamische modellen van SYSTAN/I-DEAS 6.0 naar Matlab 4.0", Philips CFT report, 1994
- HUR65 Hurty, W.C., "Dynamic Analysis of Structural Systems Using Component Modes", AIAA Journal, 1965, Vol.3, No.4, pp.678-685
- JON93 Jong, G.J. de, "Dynamische Berekeningen aan de OD100", Philips-CFT report CTB 545.93.3007, 1993
- KOS73 Koster, M.P., "Vibration of Cam Mechanisms and their Consequences on the Design", PhD-thesis, 1973, N.V. Philips' Gloeilampenfabrieken, Eindhoven, Netherlands
- MIL94 Mil, W.C.M. van, et al., "Concurrent Engineering Handboek", Philips-CFT, internal report CTR598-94-0115.
- RAN88/1 Rankers, A.M., "Analyse van het dynamisch gedrag van de CDM-8.1 paaltjes aktuator", Philips-CFT, dept. Mechanics&Mechanisms, report 880211-26, 1988
- RAN88/2 Rankers, A.M., "Dynamic Analysis CDM-8.1 (first SYSTAN model)", Philips-CFT, dept. Mechanics&Mechanisms, report 880421-27, 1988
- RAN88/3 Rankers, A.M., "Analytische Toepassing van Modal Analysis bij de bestudering van het dynamisch gedrag van constructies en mechanismen", Philips-CFT, dept. Mechanics&Mechanisms, internal report, 1988

- RAN88/4 Rankers, A.M. and Eijk, J. van, "(Non)-Linear Simulation of Servo Systems based on a modal description of the dynamics", 13th Int. Seminar on Modal Analysis. Leuven, 1988
- RAN88/5 Rankers, A.M., "Link from SYSTAN to PHILPAC+", Philips-CFT, dept. Mechanic&Mechanisms, internal note, 1988
- RAN94 Rankers, A.M. and Eijk, J. van, "The influence of reaction forces on the behaviour of high-performance motion systems", Second Int. Conference on Motion and Vibration Control, Tokyo, 1994
- ROG70 Rogers, L.C., "Derivatives of Eigenvalues and Eigenvectors", AIAA Journal, 1970, Vol.8, No.5, pp.943-944
- RUB75 Rubin, S., "Improved component mode representation for structural dynamic analysis", AIAA Journal, 1975, Vol.13, No.8 , pp.995-1006
- SCH88/1 Schrama, H.W., "ANSYS analyse van het dynamisch gedrag van de CDM-8.1 2-D actuator", Philips-CFT, dept. Mechanics & Mechanisms, internal report 880714-28, 1988
- SCH88/2 Schrama, H.W., "ANSYS analyse van varianten op de CDM-8.1 2D-actuator", Philips-CFT, dept. Mechanics&Mechanisms, internal report 880714-29, 1988
- SET89 Seto, K., Kajiwara, I. and Nagamatsu, A., "Design of an optical servo system using a structural optimization method in which controllability is taken into account", 66th Symposium of the Japan Society of Mechanical Engineers, 1989
- SET92 Seto, K., and Mitsuta, S., "A new method for making a reduced order model of flexible structures using unobservability and uncontrollability and its application in vibration control", 1st Int. Conf. on Vibration and Motion Control, Yokohama, Japan, 1992
- STE96 Steenbakkers, M.H.A., "Performance improvement of the FAMMDD - a second encoder for the inner axis", Philips-CFT report CTB 595-96-3036, 1996
- VEL89 Velthuizen, B.H.A., "Substructuring (theory and evaluation of component mode synthesis)", Master Thesis, University Twente, 1989
- WAN81 Wang, Z.W., "Accuracy of constraint mode methods in component mode synthesis techniques", Master Thesis, Institute of Sound and Vibration Research, 1981
- WAN94 Wang, B.P., "Eigensolution Derivatives for Arbitrarily Normalized Modes", Shock and Vibration, 1994, Vol. 1, No. 4, pp. 395-399

- WEI88 Wei, J.J.C., Q. Zang, R.J. Allemang and M.L. Wei, "Correction of finite element models via selected physical parameters", 13th Int. Seminar on Modal Analysis, Leuven, Belgium, 1988
- WOR94 Wortelboer, P.M.R., "Frequency-weighted Balanced Reduction of Closed-loop Mechanical Servo-systems - Theory and Tools", PhD-thesis, Delft University of Technology, 1994

Notation

This section explains most of the notation that is used throughout the thesis. To increase the traceability of symbols often applied, some symbols that are only used (and explained) very locally in the text have been left out.

General Notations

A	matrix
a	column vector
a	scalar
\dot{a}	first time derivative of a
\ddot{a}	second time derivative of a

Specific Notations

M	mass matrix
K	stiffness matrix
Φ	matrix containing the eigenvectors ϕ_i as columns
ϕ_i	eigenvector or mode i
ϕ_{ik}	element k of mode vector ϕ_i corresponding to physical DOF x_k
ω_i	natural frequency
m_i	modal mass of mode i
k_i	modal stiffness of mode i
d_i	modal damping of mode i
β_i	relative modal damping of mode i
$m_{\text{eff},ik}$	effective modal mass of mode i corresponding to physical DOF x_k

$k_{\text{eff},ik}$	effective modal stiffness of mode i corresponding to dof x_k
m_k	physical mass
k_k	physical stiffness
\mathbf{x}	vector of physical DOF
\mathbf{q}	vector of modal DOF
\mathbf{f}	force vector
x_k	physical DOF k
x_{ki}	contribution of mode i to the motion of physical DOF k
q_i	modal DOF i
f_k	force component on physical DOF k

Abbreviations

COG	Centre of Gravity
FRF	Frequency Response Function
DOF	Degree(s) of Freedom
FE	Finite Element

Appendix

A. Effect of Frame Dynamics on Settling Time

In the main part of this thesis, attention was mainly focused on stability related issues. However, one of the three major dynamic effects, namely the motion of a frame due to the reaction force induced during a set-point motion (Fig.A.1), manifests itself also very pronounced in the time domain [RAN94].

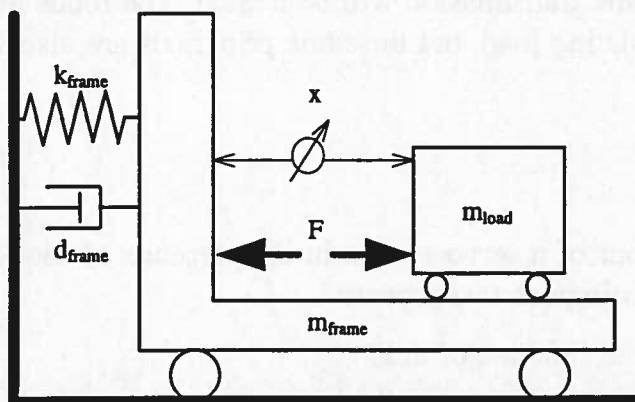


Figure A.1 Frame with limited mass and stiffness that is excited by the servo force

In a wide range of applications (even if the frame resonance is practically negligible in the mechanical transfer function due to its large relative mass), the time-domain performance is seriously limited by the effect of the vibrating frame that is excited by the servo system itself. Briefly, the phenomenon can be described as follows: The servo forces are applied on the load in order to generate a desired motion; however, the same forces also act on the frame and cause it to vibrate. At the end of the prescribed motion, the frame vibration has not ended (frame damping is limited) and the load has to "follow" the motion of the frame. The acceleration force required to do so acts as a disturbance on the servo system and inevitably introduces a positional error (Fig.A.2).

On the basis of the expected frame properties, servo bandwidth, and set-point acceleration forces, first estimates of the resulting servo error after a set-point motion can be made during the design phase. If the resulting servo error is unacceptable, several measures can be taken to reduce or eliminate this undesired phenomenon.

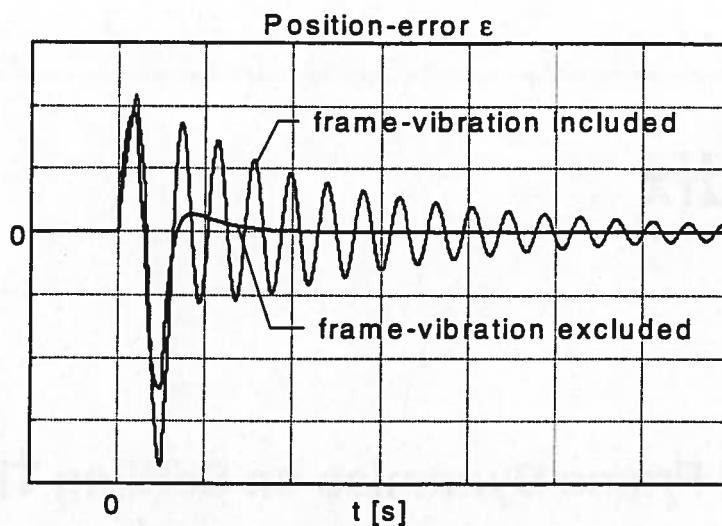


Figure A.2 Influence of frame motion on set-point response

The effect of frame vibrations becomes especially important in so-called “direct drive” systems, so some of the striking differences between this type of motor and the classical system using a motor and transmission will be treated. The focus in this appendix is on systems with a translating load, but the same principals are also valid for rotating systems.

A.1. Engineering Estimation

When analysing the settling behaviour of a servo system in the presence of set-point-induced frame vibrations, one can distinguish two aspects:

- magnitude of frame vibration as function of time
- sensitivity of the servo position to frame vibrations

Under the assumption that the presence of the servo loop does not affect the frame motion (the validity of this assumption will be discussed later), it is possible to approximate the time response of the frame motion and then derive the amount of servo error from the sensitivity plot.

The frame motion after completion of a set-point profile may be approximated by the envelope of the resulting frame displacement :

$$x_{\text{frame,env}}(t) = x_{\text{frame,0}} * e^{-\beta_{\text{frame}} \omega_{\text{frame}} t} \quad (\text{A.1})$$

in which $x_{\text{frame,0}}$ represents the amplitude of the sinusoidal motion of the frame in the absence of damping, and ω_{frame} and β_{frame} represent the natural frequency and relative damping of the frame¹.

¹ $\omega_{\text{frame}} = \sqrt{\frac{k_{\text{frame}}}{m_{\text{frame}}}}$ $\beta_{\text{frame}} = \frac{d_{\text{frame}}}{2\sqrt{m_{\text{frame}} k_{\text{frame}}}}$

The engineering estimation will be based on so-called "parabolic" or second order set-point profiles, which are characterised by three phases (1. constant acceleration, 2. constant velocity, 3. constant deceleration). In the calculation of the amount of frame vibration, one has to distinguish between two categories of set-points :

- "long" set-points with respect to the period of vibration of the frame (Fig.A.3a)
- "short" set-points with respect to the period of vibration of the frame (Fig.A.3b)

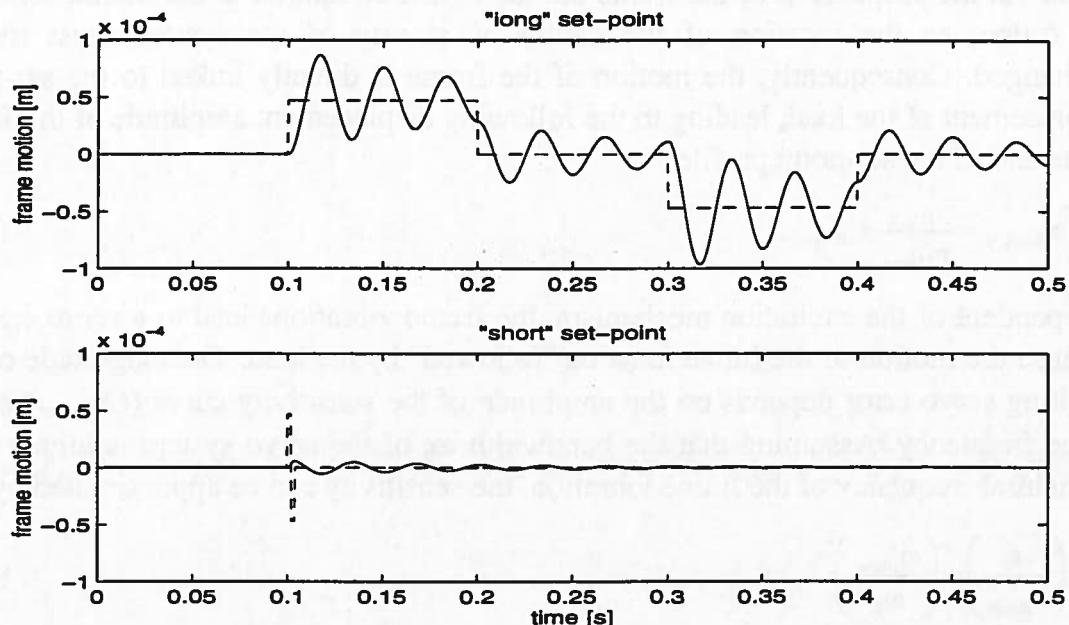


Figure A.3 Frame motion resulting from a. "long" and b. "short" parabolic set-point profile (dotted line : $m_{load} \cdot a_{set-point} / k_{frame}$)

Long Set-point Duration

In the case of a long parabolic set point, each of the four force transitions induces a frame vibration, which equals that of a force step $F = m_{load} \cdot a_{set-point}$, resulting in a displacement amplitude of the frame that can be derived from the set-point acceleration force and the frame stiffness (the mass of the frame has no influence in this load situation) :

$$x_{frame,0} = \frac{m_{load} a_{set-point}}{k_{frame}} \quad (A.2)$$

The amount of damping and the duration of the constant acceleration part determines whether the effect of each transition has already vanished at the next transition. Furthermore, depending on transition moments and the period of vibration of the frame, interference of the four transitions might occur, such that the individual effects are completely cancelled or accumulated. This effect introduces an extra multiplication factor γ that ranges from 0-4, which leads to:

$$x_{frame,0} = \gamma \frac{m_{load} a_{set-point}}{k_{frame}} ; \quad (0 \leq \gamma \leq 4) \quad (A.3)$$

Short Set-point Duration

In the case of a relatively short set-point duration, the previous approach leads to an over-estimation of the amount of frame motion. In contrast to the previous situation, the amplitude of the frame displacement is now completely dominated by the inertia forces of the frame.

During the short duration of the set-point profile, the forces that are transmitted to the world via the suspension of the frame are negligible compared to the inertia forces of the frame, so the location of the centre of gravity of the system must remain unchanged. Consequently, the motion of the frame is directly linked to the set-point displacement of the load, leading to the following displacement amplitude of the frame at the end of the set-point profile :

$$x_{\text{frame},o} = \frac{m_{\text{load}}}{m_{\text{frame}}} x_{\text{set-point}} \quad (\text{A.4})$$

Independent of the excitation mechanism, the frame vibrations lead to a servo error ε , because the motion of the frame must be "followed" by the load. The magnitude of the resulting servo error depends on the amplitude of the sensitivity curve ($\varepsilon/x_{\text{frame}}$) at the frame frequency. Assuming that the bandwidth ω_b of the servo system is higher than the natural frequency of the frame vibration, the sensitivity can be approximated by :

$$\left(\frac{\varepsilon}{x_{\text{frame}}} \right) \approx \left(\frac{\omega_{\text{frame}}}{\omega_b} \right)^2 \quad (\text{A.5})$$

Combination of the previous results with this approximation leads to the following engineering estimates for the envelope ε_{env} of the resulting servo error ε :

$$\varepsilon_{\text{env}}(t) = x_{\text{frame},o} * \left(\frac{\omega_{\text{frame}}}{\omega_b} \right)^2 * e^{-\beta_{\text{frame}} \omega_{\text{frame}} t}$$

long set-point : $x_{\text{frame},o} = \gamma \frac{m_{\text{load}} a_{\text{set-point}}}{k_{\text{frame}}} ; \quad (0 \leq \gamma \leq 4)$ (A.6)

short set-point : $x_{\text{frame},o} = \frac{m_{\text{load}}}{m_{\text{frame}}} x_{\text{set-point}}$

In the beginning of this section it was stated that this approximation is based on the assumption that the frame motion is not influenced by the control loop, and acts as an independent disturbance source. Recalling the analogy between a PD controller and a mechanical spring/damper combination (see Section 2.1), and applying it to the current phenomenon, leads to Fig.A.4, which indicates that there is a certain amount of bi-directional interaction between the load and the frame.

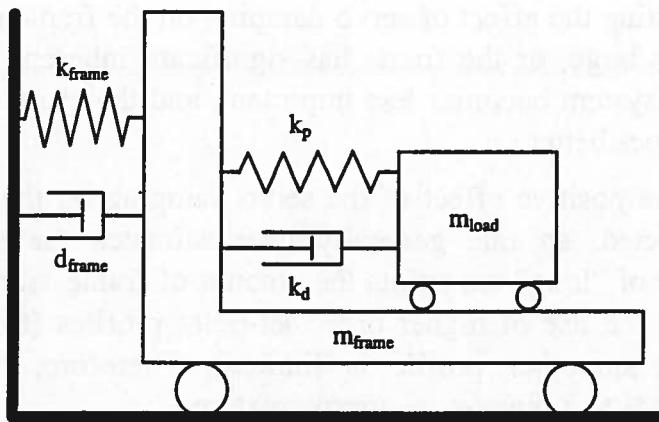


Figure A.4 Mechanical interpretation of control loop with servo stiffness k_p and servo damping k_d

Intuitively, one feels that the influence of the control loop on the frame motion is limited as long as the mass of the frame is significantly larger than that of the load². However, even at a mass ratio $m_{frame}/m_{load}=10$, a large influence can potentially be observed due to the effect that in mechanical engineering is known as “tuned damper” [HAR56]. If the bandwidth of the servo control-loop is chosen close to the natural frequency of the frame, the servo damping k_d can contribute significantly to the damping of frame motion. Furthermore, if the servo bandwidth is very high compared to the natural frequency of the frame, the additional mass of the load lowers the frame frequency in the closed-loop situation.

These effects are completely neglected in the previously presented approximations, so an investigation into the validity of these approximations has been done [BEU96]. The results of this study are roughly summarised in Table A.1, which shows the order of magnitude of *potential* deviations from the estimated settling time compared to the results of a simulation. The maximum deviations are found if the bandwidth is chosen close to the original frame frequency, because then the effect of the “tuned damping” plays an important role.

Table A.1 Potential deviation of the estimated settling times from the exact settling time

	$m_{frame}/m_{load} = 100$	$m_{frame}/m_{load} = 10$	$m_{frame}/m_{load} = 1$
$\beta_{frame}=10\%$	O(1%)	O(10%)	O(100%)
$\beta_{frame}=1\%$	O(10%)	O(100%)	>O(100%)

The general trend that can be observed in Table A.1 is that the quality of the approximation becomes better for higher mass ratios, and also for higher values of the frame damping. This trend is quite understandable, because the deviations are

² In that case the closed-loop poles corresponding to the frame motion do not differ significantly from the open-loop poles.

mainly caused by neglecting the effect of servo damping on the frame motion. If the frame mass is relatively large, or the frame has significant inherent damping, the influence of the control system becomes less important, and therefore the quality of the approximation becomes better.

In the approximation, the positive effect of the servo damping on the decay of the frame motion is neglected, so one generally over-estimates the settling time. Furthermore, in the case of "long" set-points the amount of frame vibrations can be significantly reduced by the use of higher order set-point profiles (for "short" set-points the effect of a smoother profile is limited). Therefore, the presented engineering approach leads to a worst-case approximation.

A.2. "Direct Drive" versus "Indirect Drive"

In the previous discussion, a direct-drive system without transmission was used. These systems are gaining popularity because of several advantages, such as ease of maintenance, little clearance and friction (no transmission), fewer parts, and small volume. These advantages are valid arguments for choosing a direct-drive system, but one should also be aware of the disadvantages. One is closely related to the way that servo systems react to frame motion (either self-induced during a set-point motion or resulting from external excitation, such as floor vibrations).

Figure A.5 shows a schematic representation of both a direct drive system and an indirect ball-screw driven system. Since both systems are designed with the same set-point profile and performance specification in mind, the bandwidth of both systems is identical. However, an interesting difference in actual performance can be observed when the machine frame is moving.

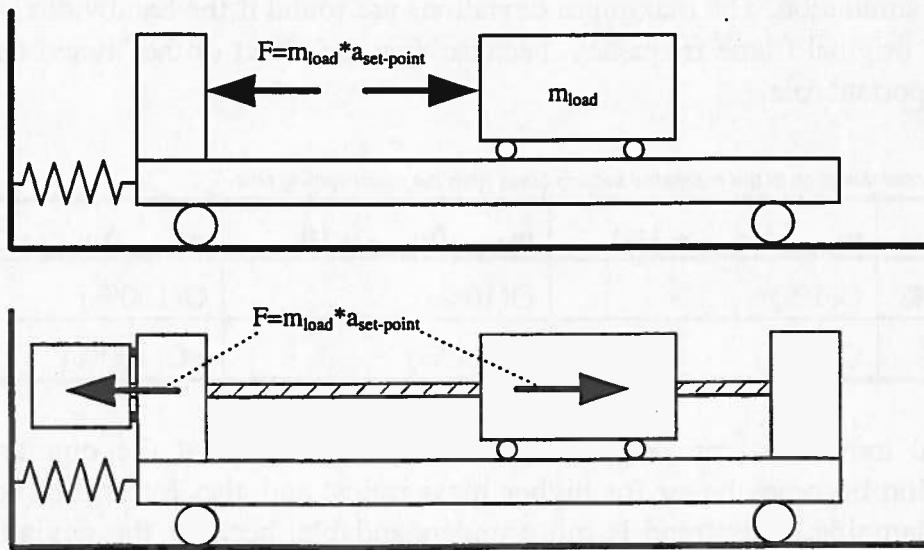


Figure A.5 Schematic representation of "direct drive" and "indirect drive"

Assuming the same motion for the load, the acceleration forces $m_{load} * a_{set-point}$ that act on the frame are identical for each of the two configurations and thus the resulting frame accelerations after a set-point are identical. Therefore, each of the systems is loaded by the same acceleration force $m_{load} * a_{frame}$. The relative displacement between load and frame that results from this disturbance force is entirely determined by the "equivalent stiffness" k_{eq} , at the load (assuming a stationary frame), which results from the servo stiffness of the position loop.

In the case of a direct-drive system the equivalent stiffness at the load that corresponds to a position bandwidth³ ω_b is given by :

$$k_{eq,direct} = k_p \approx (\omega_b)^2 * m_{load} \quad (A.7)$$

In case of the indirect-drive system with transmission ratio i from motor rotation to load translation, the total equivalent inertia m_{eq} at the load reads:

$$m_{eq,indirect} = m_{load} + J_{motor} / i^2 \quad (A.8)$$

and consequently the equivalent stiffness at the load that corresponds to a position bandwidth ω_b equals:

$$k_{eq,indirect} \approx (\omega_b)^2 * m_{eq,indirect} = (\omega_b)^2 (m_{load} + J_{motor}/i^2) \quad (A.9)$$

From equations (A.7) and (A.9) one can conclude that under the assumption of equal bandwidth, the equivalent stiffness at the load is higher in the case of an indirect drive system than a direct-drive system. This is basically caused by the increased inertia of the system. As a result of this higher equivalent stiffness an indirect drive configuration will suffer less from the effect of frame vibrations.

According to equation (A.9), decreasing the transmission ratio i (very small pitch of ball-screw) leads to an increased effective stiffness at the load. This effect is of course limited, because for very low values of the transmission ratio, the stiffness at the load is dominated by the flexibility of the transmission itself and its support. It is also interesting to note that in the case of an "inertial match"⁴ between load and motor, the effect of frame vibrations is exactly 50% of the comparable direct drive system with equal bandwidth.

³ Only the proportional gain of the controller is considered here.

⁴ Assuming that the load and the motor (and therefore the maximum motor torque) are already defined and that the transmission ratio i can still be chosen, one can maximise the acceleration of the load by selecting the transmission ratio i such that $m_{load}=J/i^2$.

A.3. Reduction of the Effect of Reaction Forces

Fundamentally, there are three ways of reducing the negative effects of the reaction force on the set-point response :

1. less set-point excitation of the frame
2. less motion of the frame, given a certain set-point excitation
3. less servo error, given a certain frame vibration

A.3.1. Less excitation of the frame

Reduction or elimination of the reaction force that is normally applied on the frame during the acceleration/deceleration phase of a set-point profile can be achieved by:

- Generating the driving force between the load and the “fixed world” instead of the frame. In many applications, however, the limited allowable motion between the two parts of the motor, in combination with the large excursion of the machine frame (if excited by other disturbances), is an important disadvantage.
- Compensating the driving force by simultaneously applying a counteracting force through external frame-motion compensation (Fig.A.6a). Effectively, this is equivalent to the first approach. However the use of a second motor, without tight requirements with respect to positioning accuracy, gives more freedom of design.
- Use of a counter mass (Fig.A.6b) (an approach that is not often seen in equipment). In this approach the reaction force is applied to a counter mass that moves in the opposite direction to the main mass instead of applying the reaction force to the frame.

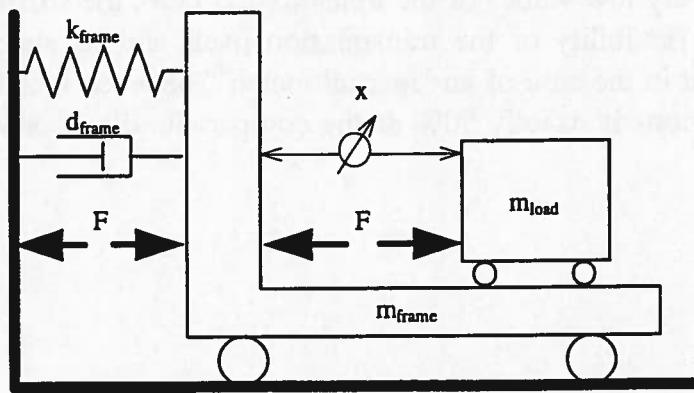


Figure A.6a Reduced frame excitation via simultaneous application of a compensation force on the frame

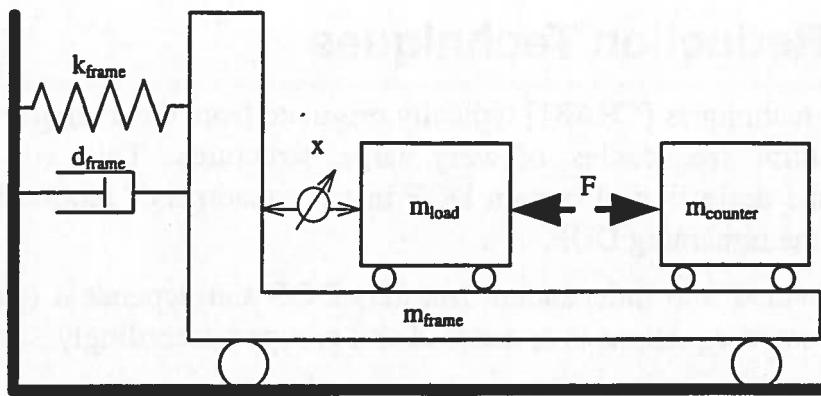


Figure A.6b Reduced frame excitation via the use of a counter mass

A.3.2. Reduced motion of the frame

If the excitation of the frame cannot be tackled sufficiently, one might try to reduce the motion of the frame by one of the following:

- more (active) damping
- higher mass of frame (effective for short set-points)
- higher stiffness of frame (effective for long set-points)

A.3.3. Increased disturbance rejection of the servo system

Finally, given a certain amount of frame motion, the effect of this motion on the servo error can be reduced by:

- higher equivalent stiffness at the load, which can be achieved by increasing the bandwidth of the servo control, by decreasing the transmission ratio in an indirect-drive system, or by changing from a direct-drive concept to an indirect-drive concept.
- application of so-called "frame acceleration compensation". In this technique, the acceleration of the frame is measured with an accelerometer and a compensating force $m_{load} * a_{frame}$ is applied to the load such that no servo error needs to be built up to provide the necessary acceleration force. With this approach, improvements of a factor 5-20 can be achieved.

B. Static Reduction Techniques

Static reduction techniques [CRA81] typically originate from the attempt to calculate modes and natural frequencies of very large structures. This was done by approximating the deflection of certain DOF in the equations of motion by a linear combination of the remaining DOF.

The DOF are divided into independent (master) DOF and dependent (slave) DOF, and the original set of equations is re-ordered and grouped accordingly :

$$\begin{bmatrix} \mathbf{M}_{mm} & \mathbf{M}_{ms} \\ \mathbf{M}_{sm} & \mathbf{M}_{ss} \end{bmatrix} \begin{bmatrix} \ddot{\mathbf{x}}_m \\ \ddot{\mathbf{x}}_s \end{bmatrix} + \begin{bmatrix} \mathbf{K}_{mm} & \mathbf{K}_{ms} \\ \mathbf{K}_{sm} & \mathbf{K}_{ss} \end{bmatrix} \begin{bmatrix} \mathbf{x}_m \\ \mathbf{x}_s \end{bmatrix} = \begin{bmatrix} \mathbf{f}_m \\ \mathbf{f}_s \end{bmatrix} \quad (\text{B.1})$$

with \mathbf{x}_m and \mathbf{x}_s representing the vectors of master DOF and slave DOF.

If the components of \mathbf{x}_s are approximated by a linear combination of the components of \mathbf{x}_m , then \mathbf{x}_s can be eliminated from the original set of equations of motion (B.1). This leads to the following reduced approximation of the original set of equations :

$$\mathbf{M}^* \ddot{\mathbf{x}}_m + \mathbf{K}^* \mathbf{x}_m = \mathbf{f}^* \quad (\text{B.2})$$

in which \mathbf{M}^* , \mathbf{K}^* and \mathbf{f}^* represent the new mass matrix, stiffness matrix, and force vector.

B.1. Static Condensation

Static condensation is based on the assumption that the kinetic energy corresponding to the motion of a limited number of DOF represents most of the kinetic energy of the system. This implies that the master DOF can be chosen at these locations with high mass compared to the rest of the structure, and that the terms \mathbf{M}_{ss} , \mathbf{M}_{sm} , and \mathbf{M}_{ms} are neglected. Assuming furthermore that the slave DOF are not loaded by external forces, the following set of equations is found :

$$\begin{bmatrix} \mathbf{M}_{mm} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \ddot{\mathbf{x}}_m \\ \ddot{\mathbf{x}}_s \end{bmatrix} + \begin{bmatrix} \mathbf{K}_{mm} & \mathbf{K}_{ms} \\ \mathbf{K}_{sm} & \mathbf{K}_{ss} \end{bmatrix} \begin{bmatrix} \mathbf{x}_m \\ \mathbf{x}_s \end{bmatrix} = \begin{bmatrix} \mathbf{f}_m \\ \mathbf{0} \end{bmatrix} \quad (\text{B.3})$$

The lower part of this set of equations gives the static relation between \mathbf{x}_s and \mathbf{x}_m :

$$\mathbf{x}_s = -\mathbf{K}_{ss}^{-1} \mathbf{K}_{sm} \mathbf{x}_m \quad (\text{B.4})$$

Substituting this relation in the upper part of the equations finally leads to the reduced equations of motion;

$$\mathbf{M}^* \ddot{\mathbf{x}}_m + \mathbf{K}^* \mathbf{x}_m = \mathbf{f}^* \quad (\text{B.5})$$

with

$$\mathbf{M}^* = \mathbf{M}_{mm}; \quad \mathbf{K}^* = \mathbf{K}_{mm} - \mathbf{K}_{ms} \mathbf{K}_{ss}^{-1} \mathbf{K}_{sm}; \quad \mathbf{f}^* = \mathbf{f}_m$$

B.2. Guyan Reduction

The Guyan reduction is basically a generalisation of the static condensation technique; it provides better kinetic energy conservation and therefore results in a better approximation of the dynamic properties. Instead of simply neglecting the mass corresponding to the slave DOF, the Guyan technique reduces the mass matrix via the same transformation as the stiffness matrix.

In the discussion of the static condensation, the reduced stiffness matrix could be derived directly from a substitution of equation (B.4) into the upper part of equation (B.3). However, this reduced stiffness matrix can also be regarded as the result of the following transformation, which is based on the idea of (potential) energy conservation :

$$\mathbf{K}^* = \mathbf{T}^T \mathbf{K} \mathbf{T} \quad (\text{B.6})$$

with \mathbf{K} being the original reshuffled stiffness matrix, and \mathbf{T} being the overall transformation matrix defined by:

$$\begin{vmatrix} \mathbf{x}_m \\ \mathbf{x}_s \end{vmatrix} = \mathbf{T} \begin{vmatrix} \mathbf{x}_m \\ \mathbf{x}_s \end{vmatrix} \quad \text{with } \mathbf{T} = \begin{bmatrix} \mathbf{I} \\ -\mathbf{K}_{ss}^{-1}\mathbf{K}_{sm} \end{bmatrix} \quad \text{and } \mathbf{I} = \text{unit matrix} \quad (\text{B.7})$$

It can be shown that the reduced stiffness matrix that is found via this approach is identical to the one found previously.

In contrast to the static condensation, the Guyan reduction does not simply set the terms \mathbf{M}_{ss} , \mathbf{M}_{sm} , and \mathbf{M}_{ms} equal to zero but calculates an approximate (reduced) mass matrix by applying this transformation to the mass matrix \mathbf{M} :

$$\mathbf{M}^* = \mathbf{T}^T \mathbf{M} \mathbf{T}$$

which after substitution of equation (B.7) results in :

$$\mathbf{M}^* = \mathbf{K}_{sm}^T \mathbf{K}_{ss}^{-1T} \mathbf{M}_{ss} \mathbf{K}_{ss}^{-1} \mathbf{K}_{sm} - \mathbf{K}_{sm}^T \mathbf{K}_{ss}^{-1T} \mathbf{M}_{sm} - \mathbf{M}_{ms} \mathbf{K}_{ss}^{-1} \mathbf{K}_{sm} + \mathbf{M}_{mm} \quad (\text{B.8})$$

The reduced stiffness matrix is identical to that of the static condensation :

$$\mathbf{K}^* = \mathbf{K}_{mm} - \mathbf{K}_{ms} \mathbf{K}_{ss}^{-1} \mathbf{K}_{sm} \quad (\text{B.9})$$

Example

The previous expressions, especially equation (B.8), look rather complicated, but the basic mechanism of the Guyan reduction can be illustrated very elegantly by considering its effect on a three-mass spring-system (Fig.B.1). In this example, x_1 and x_3 are chosen as master DOF, whereas x_2 is considered a slave DOF.

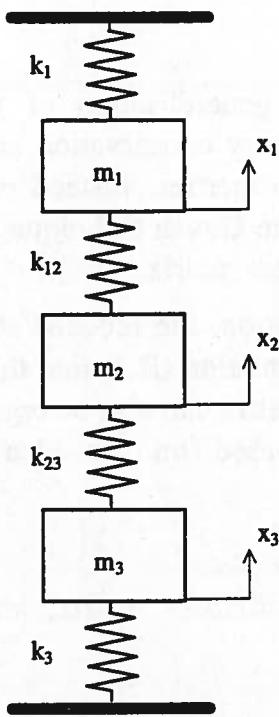


Figure B.1 Three-mass spring-system

In the absence of external forces, the original set of equations reads :

$$\begin{bmatrix} m_1 & & \ddot{x}_1 \\ & m_2 & \ddot{x}_2 \\ & & m_3 \end{bmatrix} + \begin{bmatrix} k_1 + k_{12} & -k_{12} & 0 \\ -k_{12} & k_{12} + k_{23} & -k_{23} \\ 0 & -k_{23} & k_3 + k_{23} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0 \quad (\text{B.10})$$

In correspondence with the described approach, the DOF set will be re-arranged such that the master DOF x_1 and x_3 are listed first, followed by the slave DOF x_2 :

$$\begin{bmatrix} m_1 & & \ddot{x}_1 \\ & m_3 & \ddot{x}_3 \\ & & m_2 \end{bmatrix} + \begin{bmatrix} k_1 + k_{12} & 0 & -k_{12} \\ 0 & k_3 + k_{23} & -k_{23} \\ -k_{12} & -k_{23} & k_{12} + k_{23} \end{bmatrix} \begin{bmatrix} x_1 \\ x_3 \\ x_2 \end{bmatrix} = 0 \quad (\text{B.11})$$

In this specific case, the general equation (B.4) that expresses the slave DOF in terms of the master DOF, is equal to :

$$x_2 = \frac{1}{k_{12} + k_{23}} * (k_{12} * x_1 + k_{23} * x_3) \quad (\text{B.12})$$

This equation expresses what one would have expected to find as the static deflection of x_2 , if x_1 and x_3 are prescribed. As can be seen, the static motion of x_2 is determined by x_1 and x_3 in combination with the ratio of the connecting stiffnesses and the overall stiffness at x_2 .

Substitution of this relation, and elimination of the slave DOF x_2 , finally leads to the following reduced set of equations of motion :

$$\begin{bmatrix} m_1 + \frac{k_{12}}{k_{12} + k_{23}} * m_2 & 0 & \\ 0 & m_3 + \frac{k_{23}}{k_{12} + k_{23}} * m_2 & \end{bmatrix} \begin{bmatrix} \ddot{x}_1 \\ \ddot{x}_3 \end{bmatrix} + \begin{bmatrix} k_1 + \frac{k_{12}k_{23}}{k_{12} + k_{23}} & -\frac{k_{12}k_{23}}{k_{12} + k_{23}} \\ -\frac{k_{12}k_{23}}{k_{12} + k_{23}} & k_3 + \frac{k_{12}k_{23}}{k_{12} + k_{23}} \end{bmatrix} \begin{bmatrix} x_1 \\ x_3 \end{bmatrix} = 0 \quad (\text{B.13})$$

The reduced stiffness matrix is nothing more than the stiffness that is "felt" in the master DOF. More interesting are the expressions found in the mass matrix. The extra terms that are added to m_1 and m_3 indicate that the mass corresponding to the slave DOF x_2 is distributed over the various master DOF in accordance with the stiffness that connects the slave and the master DOF.

This approach, which is physically very plausible, is illustrated in Fig.B.2. When the connection stiffnesses k_{12} and k_{23} between the slave DOF x_2 and the master DOF x_1 and x_3 are equal, the mass m_2 is distributed equally over the master DOF. However, if k_{12} is significantly larger than k_{23} , the entire mass of DOF x_2 is added to the mass of master DOF x_1 .

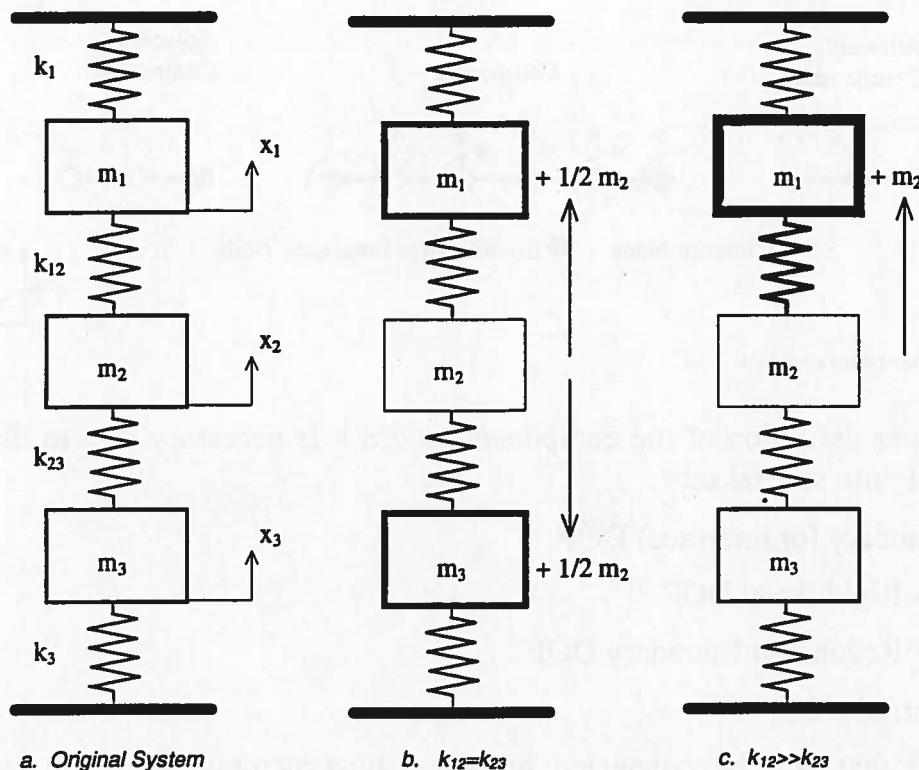


Figure B.2 Guyan reduction applied to a three-mass spring-system (x_1, x_3 are considered master DOF, x_2 is slave DOF)

C. Component Modes

In literature [CRA87, HUR65, RUB75], a wide range of component modes are proposed. As background information for the reader, the most frequently used component modes will be introduced without going into the mathematical details. This will be done using the example of a simple planar beam structure (Fig.C.1) :

- normal modes (free, fixed, hybrid, loaded)
- rigid body modes
- constraint modes
- attachment modes

The beam (component) is modelled by five elements and has six nodes, each with two DOF: translation u_y , rotation φ (u_x will not be considered). This sub-structure will be integrated into a system by connecting its end nodes to other sub-structures.

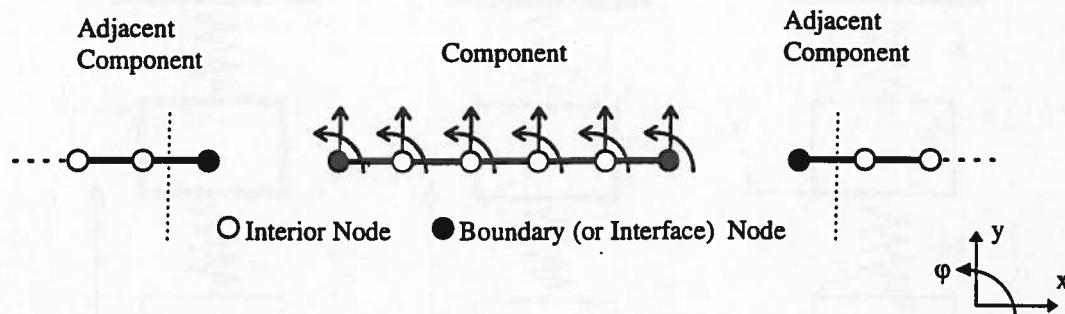


Figure C.1 Planar beam structure

For the proper definition of the component modes it is necessary first to divide the various DOF into several sets :

- Boundary (or interface) DOF
 - Rigid Body DOF
 - Redundant Boundary DOF
- Internal DOF

Those DOF that will be connected to other sub-structures when the system is assembled are called boundary (or interface) DOF (B), whereas all other DOF are internal DOF (I). The boundary DOF can further be sub-divided into a minimal set of DOF, which - when fixed - prevent the sub-structure from moving (= statically determined support). These DOF are called rigid body DOF (R), whereas the remaining boundary DOF are called redundant boundary DOF (E).

The example structure contains 12 DOF which can be divided into these categories, as indicated in Fig.C.2. As the end nodes of this beam are connected to the rest of the system, the four corresponding DOF become boundary DOF. This set of boundary

DOF is then split into two rigid body DOF (arbitrarily chosen on the left side) that are sufficient to prevent the rigid-body motion, and two redundant DOF.

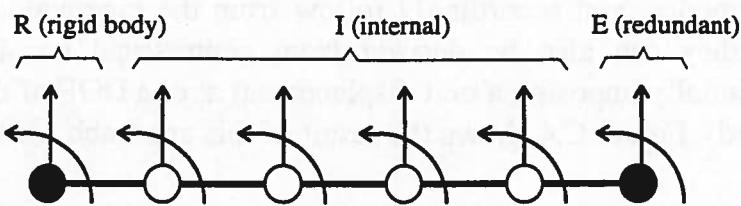


Figure C.2 Division of DOF

Normal Modes

Normal modes are the eigenvectors which result from the solution of the eigenvalue problem for the unforced vibration of the sub-structure. Depending on the type of boundary conditions, one can distinguish between:

- free interface normal modes : sub-structure completely unrestrained
- fixed interface normal modes : sub-structure fixed in all interface DOF
- hybrid interface normal modes : some interface DOF are restrained
- loaded interface normal modes : part of the mass and stiffness effects of neighbouring substructures are lumped to the boundary DOF to “simulate” the effect of the neighbouring mass and stiffness.

For the simple beam structure, the various types of normal modes are illustrated in Fig.C.3 by a sketch of the first two modes of a certain type.

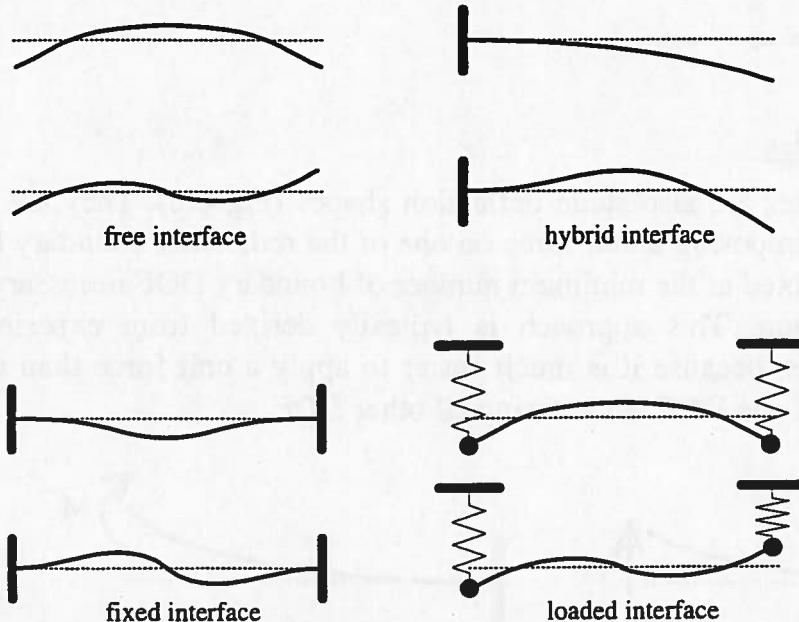


Figure C.3 Normal modes under various interface conditions

Rigid-body modes

The rigid-body modes can be considered as a special case (frequency = 0) of the free interface normal modes, and accordingly follow from the eigenvalue problem. On the other hand, they can also be derived from geometrical considerations (for example by sequentially imposing a unit displacement at one DOF of the R-set while the others are fixed). Figure C.4 shows the result of this approach for the example of the beam.



Figure C.4 Rigid body modes of the beam

Constraint Modes

Constraint modes are static deflection shapes which are obtained by successively imposing a unit deflection in one of the boundary DOF while keeping the other boundary DOF fixed. Fig.C.5 shows the resulting constraint modes of the beam structure.

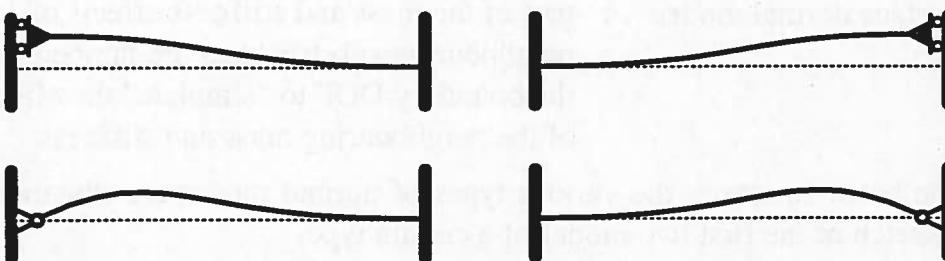


Figure C.5 Boundary constraint modes

Attachment Modes

Attachment modes are also static deflection shapes (Fig.C.6). They are determined by successively imposing a unit force on one of the redundant boundary DOF, while the structure is fixed at the minimum number of boundary DOF necessary to prevent rigid body motion. This approach is typically derived from experimental sub-structure analysis, because it is much easier to apply a unit force than to impose a unit deflection at one DOF while fixing all other DOF.



Figure C.6 Attachment modes

D. Extra Reduction of Model Size

In Section 5.3 the need for the application of component-reduction techniques, which aim at reducing the size of the component description without unacceptable consequences for the solution of the assembled system, has been discussed. However, even after application of these techniques, disk-space requirements and analysis time of the overall system can put limits on the number of design iterations that can be evaluated within a limited period of time.

Looking for ways to bypass these obstacles, it was found that the number of physical DOF retained in the component-mode description plays a significant role in the size and calculation time of the solution of the assembled system, without affecting the results of this analysis. In general, the component-mode description (see Section 5.3.1) reads :

$$\mathbf{M}_c \ddot{\mathbf{p}} + \mathbf{K}_c \mathbf{p} = \Psi^T \mathbf{f} \quad (\text{D.1})$$

$$\mathbf{x}(t) = \Psi \mathbf{p}(t) \quad (\text{D.2})$$

with

$\mathbf{x}(t)$: vector of n physical DOF

$\mathbf{p}(t)$: vector of m generalised DOF

\mathbf{M}_c : reduced component mass matrix ($m \times m$)

\mathbf{K}_c : reduced component stiffness matrix ($m \times m$)

Ψ : matrix of component modes, which describe the relation between physical and generalised DOF ($n \times m$)

n : number of physical DOF

m : number of generalised DOF

During the FE-analysis of a component, a certain element size - and linked to that a certain number of nodes and physical DOF - is required in order to properly approximate the mass and stiffness properties of a component. However, in the final component description these physical DOF x are only dependent quantities (see equation (D.2)) that have no influence on the dynamic properties of the component which are defined by the reduced mass and stiffness matrices \mathbf{M}_c , \mathbf{K}_c .

Consequently, for the further processing of a component in the system analysis stage, only the following physical DOF need to be retained :

- connection DOF : to connect the components to each other and to ground
- excitation DOF : to apply forces (servo forces, disturbing forces ...)
- response DOF : to monitor responses (servo position, frame acceleration...)

Furthermore, a sufficient number of nodes and thus physical DOF is required in order to visualise the deformation of the component.

Naturally, the number of connection, excitation, and response DOF is fixed, but the deformation of a component can easily be visualised with significantly fewer nodes (and physical DOF) than required to properly model the component properties. On the basis of geometry, mode-shapes, and some engineering judgement, it is generally quite straightforward to select those nodes that sufficiently visualise the deformation of a component.⁵

The process of reducing the number of physical DOF simply consists of deleting the superfluous data in the right hand side of equation (D.1) and in the linear equation (D.2) that relates the physical DOF to the generalised DOF. For the ease of explanation, this relation is arranged such that the superfluous physical DOF are located in \mathbf{x}_s , whereas the remaining physical DOF are combined in \mathbf{x}_r :

$$\begin{bmatrix} \mathbf{x}_r \\ \mathbf{x}_s \end{bmatrix} = \begin{bmatrix} \Psi_r \\ \Psi_s \end{bmatrix} \mathbf{p} \xrightarrow{\text{reduction}} \mathbf{x}_r = \Psi_r \mathbf{p} \quad (\text{D.3})$$

An automatic procedure for this reduction process has been implemented in the I-DEAS/SYSTAN analysis environment [GIE92/2] and is in use ever since.

As an illustration of the approach and the reduction in terms of disk space requirements and analysis time that can be achieved, the analysis of a four-axis robot system (Fig.D.1a) will be shortly discussed.

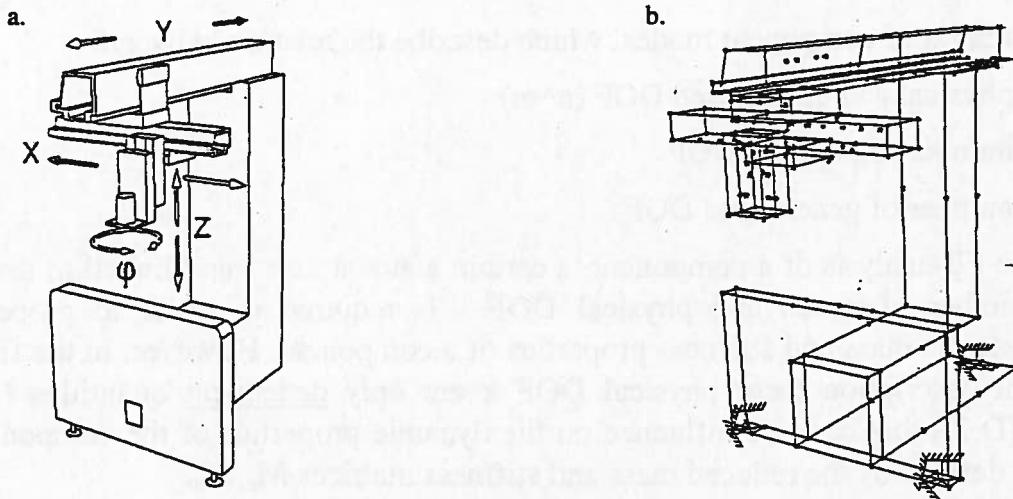


Figure D.1 Sketch of four axes robot system (a.) and final dynamic model with reduced graphical data (b.)

The dynamic model, which in its reduced form is shown in Fig.D.1b, consists of the machine frame and the beams and slides of the X, Y, and Z modules, each driven by

⁵Due to the deletion of superfluous physical DOF and nodes, elements can no longer be displayed in deformed shape. Therefore, the remaining nodes need to be connected by trace lines to represent the component.

a linear brushless motor. The φ rotation is not considered in this analysis because it adds very little extra information. In first instance, the robot frame and the Y beam are considered to dominate the overall dynamics. For that reason, these two components are modelled in FE, resulting in a component-mode description; all other components are described by their rigid-body inertia properties. The connection between the various components and the connection of the frame to ground is defined by the use of discrete springs (damping is added afterwards as modal damping).

Figure D.2 shows the original FE-meshes of the frame and the Y beam, plus the final model which only contains connection DOF, excitation DOF, response DOF, and sufficient nodes to characterise the deflection of the component.

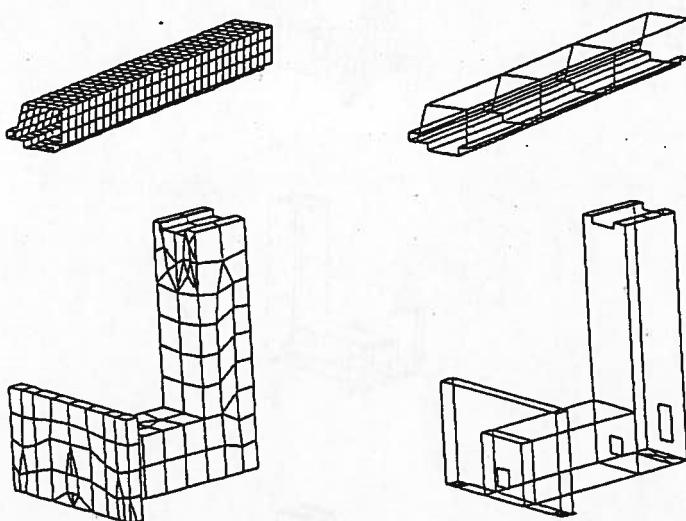


Figure D.2 Frame and Y beam (original and reduced)

The selection of visualisation nodes is a trade-off between the component size-reduction and the quality of the visualisation of deflections. As an illustration, the first four internal deformation modes of a free-free normal mode analysis of the frame structure in original and reduced state are shown in Fig.D.3.

The results of the system analysis in terms of mode-shapes and natural frequencies is not influenced by the reduction process. However, a significant reduction of disk space requirements and analysis time is achieved, as summarised in Table D.1. Apart from these improvements during the analysis phase, an extra benefit lies in the fact that the visualisation of the final system model has become much faster (for example “real time” orientation, refresh rate, animation).

The entire operation of selecting the physical DOF that need to be preserved, generating trace lines, and running the software that takes care of the reduction process, requires less than 30 minutes. Considering the enormous reduction in analysis time and the reduction of disk space makes the reduction of physical DOF a very useful modelling tool.

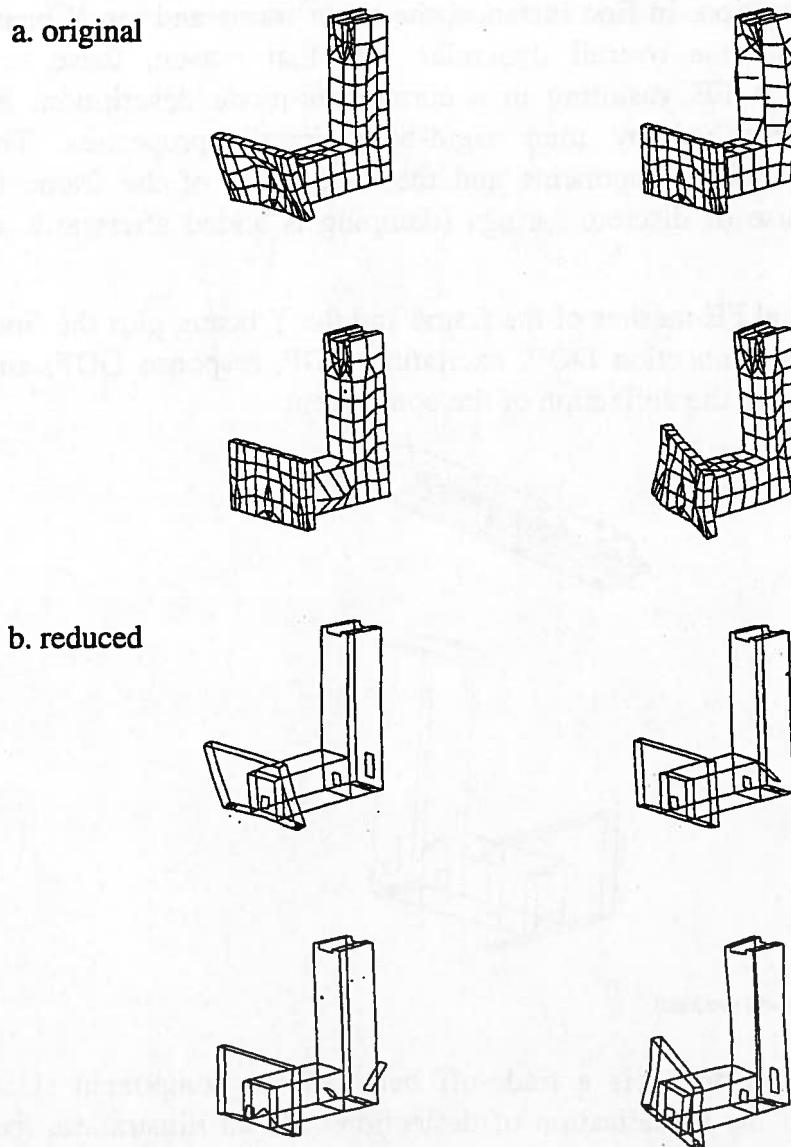


Figure D.3 First four internal deformation modes of frame (original and reduced)

Table D.1 Reduction in disk space and time

	System based on original components	System based on components with reduced graphical data
Analysis time [min]	33	10
CPU time [min]	23	6.5
Model file after solution [Mb]	43	16
Required disk space [Mb]	87	33

E. Sub-structure Coupling

For the sake of completeness, two sub-structure coupling methods will be described that are used to assemble (reduced) component models into a system model. The techniques, which both employ physical co-ordinates (coupling methods based on generalised co-ordinates are principally not different but require slightly more mathematics), are based on one of the following two approaches :

- rigid connection via elimination of DOF
- elastic connection using springs

The first method eliminates part of the combined DOF set used for coupling, thus creating a rigid connection, whereas the second method simply uses springs to create a (elastic) connection between the DOF.

These methods can best be discussed by considering two sub-structures (denoted α and β), which can each be described by a (reduced) set of equations of motion :

$$\mathbf{M}^\alpha \ddot{\mathbf{x}}^\alpha + \mathbf{K}^\alpha \mathbf{x}^\alpha = \mathbf{f}^\alpha \quad (\text{E.1a})$$

$$\mathbf{M}^\beta \ddot{\mathbf{x}}^\beta + \mathbf{K}^\beta \mathbf{x}^\beta = \mathbf{f}^\beta \quad (\text{E.1b})$$

The DOF set of each of these sub-structures can be divided into internal DOF (index I) and connection or boundary DOF (index b), which after reshuffling leads to the following set for each of the substructures :

$$\begin{bmatrix} \mathbf{M}_{bb} & \mathbf{M}_{bi} \\ \mathbf{M}_{ib} & \mathbf{M}_{ii} \end{bmatrix} \begin{bmatrix} \ddot{\mathbf{x}}_b \\ \ddot{\mathbf{x}}_i \end{bmatrix} + \begin{bmatrix} \mathbf{K}_{bb} & \mathbf{K}_{bi} \\ \mathbf{K}_{ib} & \mathbf{K}_{ii} \end{bmatrix} \begin{bmatrix} \mathbf{x}_b \\ \mathbf{x}_i \end{bmatrix} = \begin{bmatrix} \mathbf{f}_b \\ \mathbf{f}_i \end{bmatrix} \quad (\text{E.2})$$

The two sets of equations for sub-structures α and β can be combined into one overall set of equations, which is the basis for the discussion of the two coupling procedures :

$$\begin{bmatrix} \mathbf{M}_{bb}^\alpha & \mathbf{M}_{bi}^\alpha \\ \mathbf{M}_{ib}^\alpha & \mathbf{M}_{ii}^\alpha \\ \mathbf{M}_{bb}^\beta & \mathbf{M}_{bi}^\beta \\ \mathbf{M}_{ib}^\beta & \mathbf{M}_{ii}^\beta \end{bmatrix} \begin{bmatrix} \ddot{\mathbf{x}}_b^\alpha \\ \ddot{\mathbf{x}}_i^\alpha \\ \ddot{\mathbf{x}}_b^\beta \\ \ddot{\mathbf{x}}_i^\beta \end{bmatrix} + \begin{bmatrix} \mathbf{K}_{bb}^\alpha & \mathbf{K}_{bi}^\alpha \\ \mathbf{K}_{ib}^\alpha & \mathbf{K}_{ii}^\alpha \\ \mathbf{K}_{bb}^\beta & \mathbf{K}_{bi}^\beta \\ \mathbf{K}_{ib}^\beta & \mathbf{K}_{ii}^\beta \end{bmatrix} \begin{bmatrix} \mathbf{x}_b^\alpha \\ \mathbf{x}_i^\alpha \\ \mathbf{x}_b^\beta \\ \mathbf{x}_i^\beta \end{bmatrix} = \begin{bmatrix} \mathbf{f}_b^\alpha \\ \mathbf{f}_i^\alpha \\ \mathbf{f}_b^\beta \\ \mathbf{f}_i^\beta \end{bmatrix} \quad (\text{E.3})$$

E.1. Rigid connection

In the case of a rigid connection of the two sub-structures in the coupling DOF, the motion of the corresponding coupling DOF of both components is identical, thus :

$$\mathbf{x}_b^\alpha = \mathbf{x}_b^\beta = \mathbf{x}_b \quad (\text{E.4})$$

Substitution of this relation into the uncoupled set of equations leads to the following coupled set of equations based on rigid connection in the boundary DOF :

$$\begin{bmatrix} M_{bb}^\alpha + M_{bb}^\beta & M_{bi}^\alpha & M_{bi}^\beta \\ M_{ib}^\alpha & M_{ii}^\beta & 0 \\ M_{ib}^\beta & 0 & M_{ii}^\beta \end{bmatrix} \begin{bmatrix} \ddot{x}_b^\alpha \\ \ddot{x}_i^\alpha \\ \ddot{x}_i^\beta \end{bmatrix} + \begin{bmatrix} K_{bb}^\alpha + K_{bb}^\beta & K_{bi}^\alpha & K_{bi}^\beta \\ K_{ib}^\alpha & K_{ii}^\beta & 0 \\ K_{ib}^\beta & 0 & K_{ii}^\beta \end{bmatrix} \begin{bmatrix} x_b^\alpha \\ x_i^\alpha \\ x_i^\beta \end{bmatrix} = \begin{bmatrix} f_b^\alpha + f_b^\beta \\ f_i^\alpha \\ f_i^\beta \end{bmatrix} \quad (E.5)$$

E.2. Elastic connection using springs

The second technique connects the boundary DOF through the addition of a spring element (more general : a stiffness matrix) between the corresponding boundary DOF of the two sub-structures. This approach offers the possibility to add flexibility between the connecting DOF to account for various unwanted intermediate flexibilities (for example Hertz contact flexibility, limited stiffness of a glued or bolted connection) or functional flexibility (for example rubber vibration isolator).

Considering only the boundary DOF of both sub-structures, the connection stiffness matrix reads :

$$\begin{bmatrix} C_{bb} & -C_{bb} \\ -C_{bb} & C_{bb} \end{bmatrix} \quad (E.6)$$

where C_{bb} is a stiffness matrix working on the boundary DOF.

Substitution of this relation in the uncoupled set of equations leads to the following set of coupled equations:

$$\begin{bmatrix} M_{bb}^\alpha & M_{bi}^\alpha & \ddot{x}_b^\alpha \\ M_{ib}^\alpha & M_{ii}^\alpha & \ddot{x}_i^\alpha \\ M_{bb}^\beta & M_{bi}^\beta & \ddot{x}_b^\beta \\ M_{ib}^\beta & M_{ii}^\beta & \ddot{x}_i^\beta \end{bmatrix} + \begin{bmatrix} K_{bb}^\alpha + C_{bb} & K_{bi}^\alpha & -C_{bb} & 0 \\ K_{ib}^\alpha & K_{ii}^\alpha & 0 & 0 \\ -C_{bb} & 0 & K_{bb}^\beta + C_{bb} & K_{bi}^\beta \\ 0 & 0 & K_{ib}^\beta & K_{ii}^\beta \end{bmatrix} \begin{bmatrix} x_b^\alpha \\ x_i^\alpha \\ x_b^\beta \\ x_i^\beta \end{bmatrix} = \begin{bmatrix} f_b^\alpha \\ f_i^\alpha \\ f_b^\beta \\ f_i^\beta \end{bmatrix} \quad (E.7)$$

Which of the two approaches should be used depends very much on the application. The first approach can only be used to create an infinitely stiff, rigid connection, whereas the latter approach is more appropriate where connection flexibility needs to be taken into account. Although it is possible to approximate a rigid connection by using very high values for the connection stiffness (matrix) C_{bb} compared to the internal stiffness of the structure, it is not advisable to do so because singularity problems are very likely to occur. In practice, one obtains good results as long as physically realistic connection stiffnesses are chosen.

F. User-Defined DOF

Often it is not the response of an individual DOF one is interested in, but rather some quantity such as the relative displacement between a printed circuit board and the placement nozzle of a component-mounter, or the position of the laser spot relative to the track on a compact disc, that can be defined as a linear combination of individual DOF responses. The relation for such a quantity, which is often referred to as "user DOF", can be defined as :

$$x_{\text{user}} = b_1 * x_1 + b_2 * x_2 + \dots + b_n * x_n \quad (\text{F.1})$$

The definition of such a user DOF is a very practical feature of many analysis programs, because it eliminates the cumbersome "construction" of time and frequency responses as a combination of individual responses.

This appendix has been added to stretch the importance of a correct definition of user DOF, because experience teaches that despite its apparent triviality many mistakes are often introduced in the definition of such a user DOF. This is often caused by a lack of attention in combination with unawareness of the detrimental effects that can be introduced.

Obviously, the formula for a user DOF must include the motion of all relevant components. But apart from that it is essential that the expression for a user DOF, if it measures some sort of internal deformation of the system, is invariant to real rigid-body motions. As a mental experiment, one should eliminate all supports (fixations to ground) of the system, and give the system an arbitrary orientation. In the absence of gravity effects, the resulting value of the user DOF should then be zero, because there is no internal deformation.

To illustrate the definition of user DOF, one can look at a model of a machine consisting of a frame plus suspension, and a slide. Figure F.1 shows a one-dimensional model, in which case the definition of the sensor DOF (= user DOF) is straightforward :

$$x_{\text{sensor}} = x_{\text{slide}} - x_{\text{frame}} \quad (\text{F.2})$$

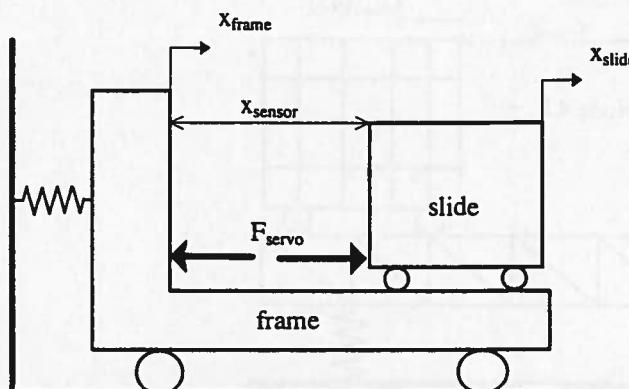


Figure F.1 Definition of user DOF in the case of a 1D model ($x_{\text{sensor}} = x_{\text{slide}} - x_{\text{frame}}$)

In the case of a 2D model, which is defined by the position and orientation of the centre of gravity of each of the two components, more attention is required to properly define the sensor DOF. If one simply copied the sensor definition of the 1D model, any static rigid-body rotation γ of the entire system would result in a sensor output equal to $\gamma \cdot h$, if h denotes the distance between the two COGs. Such a definition would lead to incorrect results. The correct formulation should include the rotational effects as well, and reads :

$$x_{\text{sensor}} = (x_{\text{slide}} - \varphi_{\text{slide}} * l_{\text{slide}}) - (x_{\text{frame}} - \varphi_{\text{frame}} * l_{\text{frame}}) \quad (\text{F.3})$$

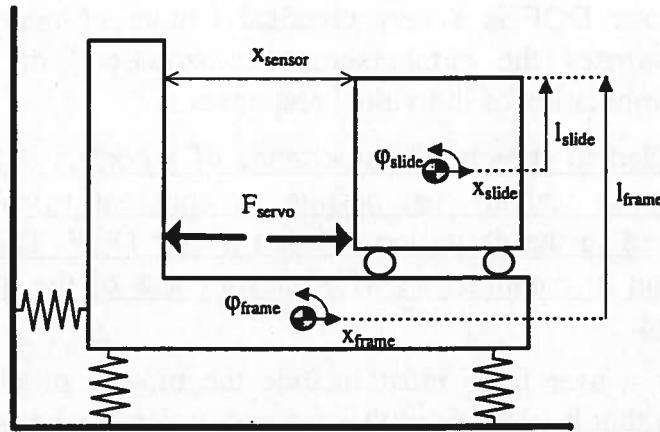


Figure F.2 Definition of user DOF in the case of a 2D model ($x_{\text{sensor}} = (x_{\text{slide}} - \varphi_{\text{slide}} * l_{\text{slide}}) - (x_{\text{frame}} - \varphi_{\text{frame}} * l_{\text{frame}})$)

In the case of a detailed FE model of the system, the nodes created by the mesh generator (Fig.F.3) will generally - unless special precautions are taken - not be located such that a proper sensor DOF can be easily defined. Users are then sometimes tempted to use nodal displacements that are close to the desired location (in the example one might define a user DOF $x_{\text{sensor}} = x_{\text{frame}}(13) - x_{\text{slide}}(47)$), which again leads to incorrect results.

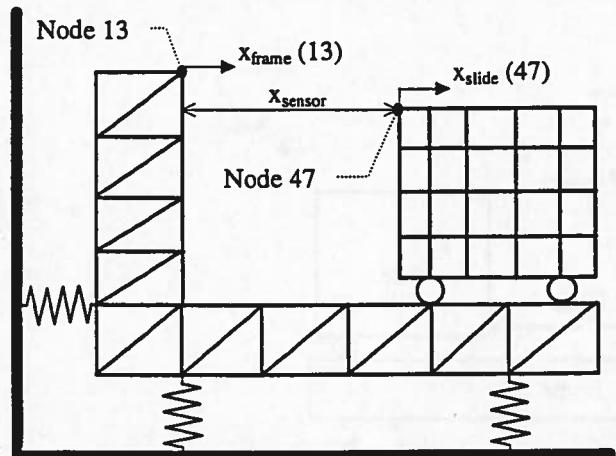


Figure F.3 Hypothetical mesh of a FE-model, from which it is very difficult to define a proper user DOF x_{sensor}

Depending on the severity of the mismatch, and its contribution compared to other effects, this error can be huge or negligible, but one has to be aware of its existence.

A good working method is to avoid any discussion about potential influences due to this effect by simply forcing the mesh generator to create nodes at predefined locations, moving nodes by hand to such a location, using element interpolation functions or by compensating these mismatches by adding extra terms in the user DOF definition. Alternatively, one could also attach an infinitely stiff and massless beam to one of the two components and create an artificial node at the correct location. The latter approach requires some engineering judgement in the selection of the beam properties. To avoid numerical problems, the values of the stiffness and the mass should not be too extreme, but deformation, and therefore influence on the final result, should also be avoided. A test with two different settings leading to the same overall results is probably the best guarantee for success.

Example : Wafer-stepper

The main function of a wafer-stepper is the projection of the sub-micron patterns of a reticle (comparable to a slide of a slide projector) via an attenuation lens on a silicon wafer (Fig.F.4). In the analysis of the dynamic performance of such a system it is important to know how well the reticle is projected on to the wafer. Avoiding all the optical details, this question can be answered by analysing how well the three important components, reticle, lens (optical centre), and wafer, are aligned.

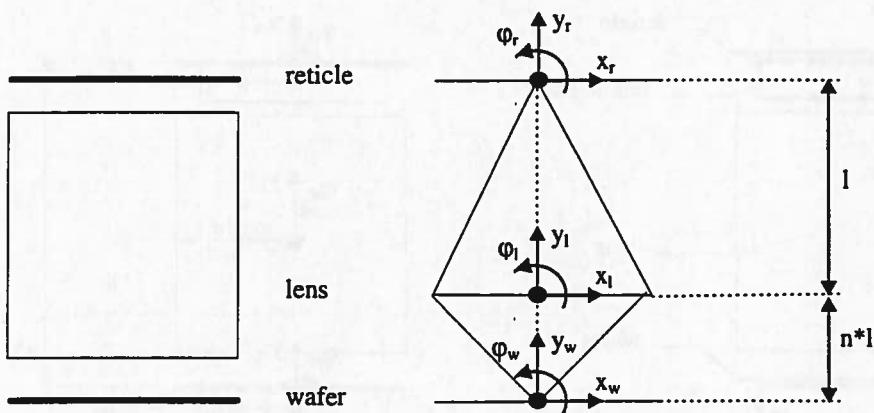


Figure F.4 Alignment of reticle, lens and wafer

The quality of the alignment can be expressed in the following user defined DOF :

$$x_{\text{error}} = x_w + n \cdot x_r - (1+n) \cdot x_l \quad (\text{F.4})$$

in which n stands for the absolute value of the lens magnification-factor.

Again, this relation can easily be checked by considering various conditions of a rigid body motion of the entire system. For each of these conditions, the formula F.4 should give a zero value⁶.

- translation in x-dir. : $x_{\text{error}} = x + n*x - (1+n)*x = 0$
- rotation φ about wafer : $x_{\text{error}} = 0 + n*(-\varphi*(1+n)*l) - (1+n)*(-\varphi*n*l) = 0$
- rotation φ about reticle : $x_{\text{error}} = \varphi*(1+n)*l + 0 + (1+n)*(0) = 0$

Equation F.4 expresses the fundamental relation of the alignment of reticle, lens and wafer. However, in the simulation model of the wafer-stepper only the wafer stage (which supports the wafer) and the reticle stage (which holds the reticle) are modelled, and not the wafer and reticle itself (Fig.F.5). Furthermore, the lens is typically modelled as one rigid body and only the displacement and rotations of the COG are available. In that case, the formula for the optical relation needs to be extended by substituting the following expressions into equation (F.4) :

$$\begin{aligned} x_w &= x_{ws} - \varphi_{ws} * l_w \\ x_r &= x_{rs} - \varphi_{rs} * l_r \\ x_l &= x_{lc} + \varphi_{lc} * l_l \end{aligned} \quad (\text{F.5})$$

Again, very much care must be taken regarding the proper distances between the various components, both in the formula and in the simulation model, in order to obtain meaningful results.

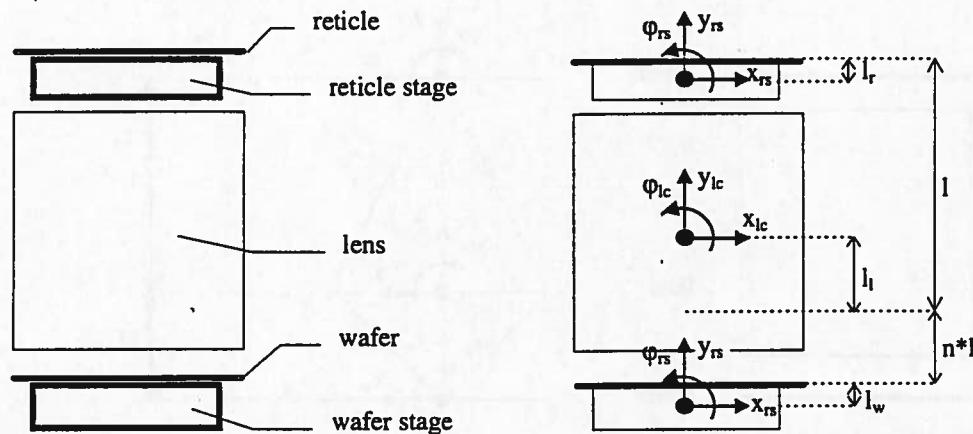


Figure F.5 Schematic representation of wafer (stage), lens and reticle (stage)

⁶A first condition for the correctness of the optical relations is that the theoretical outcome of these equations is zero for a rigid-body motion of the entire system. A second condition is the proper distance between the DOF which are used to define the optical relation. Rigid-body rotations will inevitably lead to non-zero results of the optical relation if the actual distances, lens-wafer and lens-reticle, in the simulation model differ from the values used in the optical relation.

Samenvatting

Ontwerpers van producten en machines worden geconfronteerd met steeds verder toenemende eisen ten aanzien van specificaties, zoals grotere functionaliteit, hogere snelheden en kleinere toleranties. Typische voorbeelden van deze tendens zijn de bekende Compact Disc speler en apparatuur, die bij de fabrikage van geïntegreerde schakelingen wordt ingezet. Om aan deze eisen tegemoet te komen wordt op steeds grotere schaal gebruik gemaakt van intelligente besturingen en regelsystemen. Bij de ontwikkeling van dergelijke geregelde servo positioneer-systemen is het essentieel om de dynamische eigenschappen van de mechanische constructie te beschouwen, omdat deze de volgende effecten kan bewerkstelligen :

- Mechanische resonanties kunnen de stabiliteit van de regellus in gevaar brengen en beperken derhalve de maximaal haalbare bandbreedte en de mate van storingsonderdrukking.
- Trillingen van het mechanisch systeem, die b.v. geïntroduceerd worden door de versnellingskrachten tijdens het opgelegde bewegingsprofiel, kunnen tot dynamische positiefouten leiden. In een point-to-point applicatie kan dit aanleiding zijn om aan het einde van de gewenste beweging een zekere wachttijd t.b.v. het uittrillen te introduceren. Wachttijd biedt geen soelaas in een applicatie, waarbij een baan nauwkeurig moet worden doorlopen.

Ondanks het feit dat mechanische trillingen in het algemeen vaak ingewikkeld zijn en er veel componenten van de machine bij betrokken zijn, kunnen een drietal kenmerkende probleemsoorten worden onderscheiden, die het best als volgt kunnen worden omschreven :

- actuator-flexibiliteit
- geleidingsflexibiliteit
- beperkte massa en stijfheid van het machine-frame

De basis-karakteristiek van actuator-flexibiliteit is het feit, dat de actuator zich niet als één star lichaam gedraagt t.g.v. een flexibiliteit tussen dat deel van de actuator waarop de servokrachten aangrijpen en dat deel welk nauwkeurig dient te worden gepositioneerd. De tweede categorie van dynamische verschijnselen ontstaat door geleidingsflexibiliteit. Als de lineaire actuator niet in zijn massa-middelpunt wordt aangedreven, moet een beroep worden gedaan op de (eindige) stijfheid van de

geleiding om ongewenste bewegingen te onderdrukken. De beperkte massa en stijfheid van het machine frame leidt ertoe, dat de reactiekrachten die altijd gepaard gaan met de servo-aandrijfkrachten een beweging van het machine-frame bewerkstelligen, die op zijn beurt tot positiefouten leidt.

Een veelvuldig toegepaste techniek om ingewikkelde trillingen en de onderliggende oorzaken in kaart te brengen is de "Modale Analyse" of "Modale Decompositie". Deze aanpak heeft reeds een lange traditie, zowel als experimentele maar ook als numerieke techniek. Door het "vertalen" van een ontkoppelde modale vergelijking naar een grafische representatie, die alle relevante data, zoals (effectieve) modale massa en stijfheid en de beweging van de fysische vrijheidsgraden bevat, wordt de techniek eenvoudiger te hanteren. Deze representatie leidt tot een intuitieve verbinding tussen het modale en het fysische domein en vergemakkelijkt een creatief gebruik van de modale analyse.

Resonanties van de mechanische constructie in een servosysteem kunnen leiden tot stabiliteitsproblemen van de regellus. Het is derhalve belangrijk om de frequentie-responsie-functie ($x_{\text{servo}}/F_{\text{servo}}$) te onderzoeken, die het dynamisch gedrag van het mechanische systeem karakteriseert. Uitgaande van een goed begrip van het gedrag van ontkoppelde trilvormen is het niet moeilijk om een dergelijke frequentie-responsie-functie te construeren uit de interactie van de gewenste (starre lichaam) beweging en de bijdrage van een additionele mode. Hierbij kunnen vier interactie patronen worden onderscheiden :

- -2 helling/nulpunt/pool/-2 helling
- -2 helling/pool/nulpunt/-2 helling
- -2 helling/pool/-4 helling
- -2 helling/pool/-2 helling

Het fase diagram van een -2 helling/nulpunt/pool/-2 helling karakteristiek vertoont een fase voorsprong tussen de frequentie van het nulpunt (of anti-resonantie) en die van de pool (of resonantie), terwijl de omgekeerde volgorde van nulpunt en pool bij een -2 helling/pool/nulpunt/-2 helling patroon juist voor een fase achterstand zorgt. Bij een -2 helling/pool/-4 helling gedraging draait de fase bij de frequentie van de pool van -180° naar -360° . De laatste karakteristiek, welke behoort bij een niet-minimum-fasesysteem, treedt in de normale praktijk niet op en wordt derhalve buiten de discussie gehouden. De aanpak waarbij een frequentie responsie functie wordt opgebouwd uit de beweging van de actuator als star lichaam en de bijdrage van één additionele mode kan ook met succes worden toegepast op de drie eerder genoemde dynamische basis mechanismen.

Het is niet mogelijk om een uitspraak te doen over het potentieel destabiliserende effect van elk van de genoemde karakteristieke frequentie-responsie-functies zonder de frequentie van de mechanische resonantie in relatie tot de beoogde bandbreedte

van het servosysteem te beschouwen. Het fase diagram van een typische open lus van een mechanisch servosysteem zonder resonanties (mechanica, PID-regelaar, laagdoorlaat-filter) kan in drie gebieden onderverdeeld worden, die door -180° doorgangen worden gescheiden. Laagfrequent ligt de fase onder -180° ten gevolge van de integrerende werking van de regelaar. In het midden frequente gebied (rondom de bandbreedte) ligt de fase boven -180° ten gevolge van de differentierende actie van de regelaar welke nodig is om een stabiele positielus te realiseren. Hoogfrequent zal de fase uiteindelijk weer afvallen en onder de -180° liggen door toedoen van het laagdoorlaat-filter, beperkte bandbreedte van componenten in de lus of door de beperkte bemonsteringsfrequentie van de digitale regelaar.

Met dit ideale open lus gedrag als referentie kan het destabiliserende gedrag van elk van de drie karakteristieke patronen beoordeeld worden in relatie tot het frequentiegebied waarin dit optreedt. Uiteraard hangt het daadwerkelijke optreden van instabiliteit ook af van de amplitude en damping van de additionele mode.

- Een -2 helling/nulpunt/pool/ -2 helling patroon leidt tot een fase voorsprong en kan derhalve in het laagfrequente en in het hoogfrequente gebied instabiliteit veroorzaken. In het middenfrequente gebied wordt de reeds aanwezige fase voorsprong enkel vergroot en treedt geen instabiliteit op.
- Een -2 helling/pool/nulpunt/ -2 helling combinatie heeft het tegenovergestelde effect. In het middenfrequente gebied kan bij een dergelijk gedrag instabiliteit optreden terwijl dit niet het geval is in het laagfrequente en hoogfrequente gebied.
- Indien een -2 helling/pool/ -4 helling patroon in het laagfrequente of middenfrequente gebied optreedt zal dit vrijwel altijd tot stabiliteitsproblemen leiden. Slechts in het hoogfrequente gebied kan een dergelijk patroon getolereerd worden.

Op basis van deze beschouwingen is het mogelijk om ontwerprichtlijnen op te stellen voor de mechanica van servo positionersystemen, waarbij in het bijzonder wordt gekeken naar de drie fundamentele dynamica-verschijnselen. Het uitgangspunt hierbij is het minimaliseren van de kans dat door het optreden van stabiliteitsproblemen de benodigde bandbreedte van het geregelde systeem niet kan worden gerealiseerd.

Omdat machinedynamica en de interactie met het regelsysteem een dominante rol speelt bij snelle en nauwkeurige positionersystemen, is het essentieel om deze aspecten reeds in een vroegtijdig stadium van het ontwerpproces te beschouwen. Modelvorming en simulatie kunnen hierbij een wezenlijke rol spelen, mits aan de volgende condities wordt voldaan :

- bruikbare resultaten
- snelheid

Het is voor de hand liggend, doch het wordt vaak vergeten, dat analyse-resultaten met de juiste hoeveelheid detail en nauwkeurigheid aanwezig moeten zijn op de beslissingsmomenten. Tevens is het vaak nodig om de analyse-data te vertalen naar bruikbare informatie op basis waarvan ontwerpbeslissingen genomen kunnen worden.

Aan deze voorwaarden kan slechts worden tegemoet gekomen indien het analyse-traject sterk gekoppeld is aan het ontwerptraject. Tijdens het ontwerpproces, dat normaliter start met de generatie van globale conceptideeën en uiteindelijk convergeert naar één detailontwerp, groeit de hoeveelheid informatie en detail gestaag. Het analyseproces dient derhalve op een soortgelijke top-down benadering gebaseerd te zijn. Beginnend met zeer elementaire simulatie-modellen om de conceptkeuzes te ondersteunen worden de simulatie-modellen steeds verfijnder naar gelang het ontwerpproces vordert.

In verschillende projecten, die gedurende de afgelopen jaren zijn uitgevoerd, is een stapsgewijze modelvormingsaanpak geëvolueerd, waarin de volgende fasen onderscheiden kunnen worden :

- concept-analyse
- systeem-analyse
- component-analyse

In de concept-analyse wordt de levensvatbaarheid van verscheidene concepten geëvalueerd op basis van zeer eenvoudige modellen, die uit een beperkt aantal diskrete massa's en veren bestaan. Zodra een concept gekozen is en de eerste 3D schetsen beschikbaar komen kan een systeem-analyse worden uitgevoerd op basis van een beperkt aantal 3D starre lichamen, die middels veren met elkaar verbonden zijn. In deze fase wordt een grote hoeveelheid ruimtelijke informatie toegevoegd, zoals de locatie van massamiddelpunten, verbindingsstijfheden, aandrijfkrachten en sensoren. Uiteindelijk wordt in de component-analyse de interne dynamica van belangrijke componenten onderzocht middels eindige elementen berekeningen. In gevallen, waarin een separate analyse van het dynamisch gedrag van een component onvoldoende inzicht geeft in het effect ervan op het totale systeem, kan het nodig zijn om de gedetailleerde component beschrijving te incorporeren in het systeemmodel ter vervanging van de eerdere starre modellering.

In het geval dat veel componenten van het systeem gedetaileerd gemodelleerd dienen te worden, is het niet praktisch (foutgevoelig, grote modellen, tijdrovend) om één enkel groot eindige elementen model van het totale systeem te genereren. Een techniek, die deze problemen voorkomt is de zogenaamde "Substructuring Techniek". In deze aanpak wordt het systeem onderverdeeld in sub-systemen of componenten, die separaat geanalyseerd worden. Na toepassing van een reductie techniek op elk van de component modellen, die echter de meest dominante dynamische eigenschappen van een component bewaart, worden deze

samengebouwd tot een (gereduceerd) model van het totale systeem. Hierdoor wordt de afmeting van het uiteindelijk systeemmodel aanzienlijk gereduceerd. De toegepaste reductie technieken kunnen grofweg in twee categorieën worden onderverdeeld :

- Statische Reductie Technieken
- Component Mode Technieken

Beargumenteerd wordt, waarom de component mode techniek op basis van de zo-genaamde Craig-Bampton beschrijving geprefereerd wordt.

Niet alle programmatuur die in de verschillende fasen gebruikt wordt bij de modellering van de mechanische constructie is in staat om rechtstreeks een tijd-domein simulatie uit te voeren van het complete servosysteem, inclusief de regelaar. Daardoor is het nodig om de dynamische eigenschappen van de mechanische constructie over te brengen naar andere programmatuur, waarin het regelsysteem kan worden toegevoegd. Deze overdracht is in het algemeen gebaseerd op een beperkte modale beschrijving van het mechanisch systeem, die vertaald wordt naar een toestandsmodel.

Veel van de inzichten en technieken, die in deze dissertatie zijn beschreven, zijn geëvolueerd tijdens projecten die gerelateerd zijn aan de ontwikkeling van compact disc loopwerken. De ontwikkeling van dergelijke systemen is een uitdagende bezigheid, omdat submicron nauwkeurigheid dient te worden bereikt tegen een discount prijs en onder extreme omgevingscondities (temperatuur, jogging ...). Mechanische resonanties en de invloed ervan op het regelsysteem is één van de beperkende aspekten in het ontwerp van dergelijke systemen en zal door de steeds toenemende specificaties (snelheden, opslagdichtheid) ook in de toekomst een belangrijke rol blijven spelen. Bijna alle onderwerpen, die in dit proefschrift behandeld zijn, komen aan de orde bij de bespreking van een case study, die betrekking heeft op de ontwikkeling van de CDM-8 compact disc module.

De enorme prestatie-jacht in mechatronische systemen heeft ertoe geleid, dat in de afgelopen jaren veel kennis is opgebouwd omtrent de invloed van machinedynamica in servo-positioneersystemen. De beperkte hoeveelheid gecompliceerde wiskunde in dit proefschrift toont aan, dat deze belangrijke aspekten goed kunnen worden beschreven en begrepen uitgaande van operationele kennis van basis begrippen uit de machine dynamica en de regeltechniek. De inhoud van dit proefschrift wordt derhalve beschouwd als overdraagbare kennis, die in de toekomst deel zou kunnen uitmaken van de opleiding tot mechatronisch ingenieur.

deze voorstellingen van de politieke en sociale ontwikkeling in de wereld en de verschillende landen. De voorstellingen zijn gebaseerd op de gedachten en ervaringen van de schrijver.

De voorstellingen zijn geschreven door een aantal verschillende mensen en hebben verschillende stijlen en wijzen van benadering. Ze zijn geschreven voor verschillende doeleinden en hebben verschillende doelgroepen.

De voorstellingen zijn geschreven door verschillende mensen en hebben verschillende stijlen en wijzen van benadering. Ze zijn geschreven voor verschillende doeleinden en hebben verschillende doelgroepen.

De voorstellingen zijn geschreven door verschillende mensen en hebben verschillende stijlen en wijzen van benadering. Ze zijn geschreven voor verschillende doeleinden en hebben verschillende doelgroepen.

De voorstellingen zijn geschreven door verschillende mensen en hebben verschillende stijlen en wijzen van benadering. Ze zijn geschreven voor verschillende doeleinden en hebben verschillende doelgroepen.

De voorstellingen zijn geschreven door verschillende mensen en hebben verschillende stijlen en wijzen van benadering. Ze zijn geschreven voor verschillende doeleinden en hebben verschillende doelgroepen.

De voorstellingen zijn geschreven door verschillende mensen en hebben verschillende stijlen en wijzen van benadering. Ze zijn geschreven voor verschillende doeleinden en hebben verschillende doelgroepen.

De voorstellingen zijn geschreven door verschillende mensen en hebben verschillende stijlen en wijzen van benadering. Ze zijn geschreven voor verschillende doeleinden en hebben verschillende doelgroepen.

De voorstellingen zijn geschreven door verschillende mensen en hebben verschillende stijlen en wijzen van benadering. Ze zijn geschreven voor verschillende doeleinden en hebben verschillende doelgroepen.

De voorstellingen zijn geschreven door verschillende mensen en hebben verschillende stijlen en wijzen van benadering. Ze zijn geschreven voor verschillende doeleinden en hebben verschillende doelgroepen.

De voorstellingen zijn geschreven door verschillende mensen en hebben verschillende stijlen en wijzen van benadering. Ze zijn geschreven voor verschillende doeleinden en hebben verschillende doelgroepen.

De voorstellingen zijn geschreven door verschillende mensen en hebben verschillende stijlen en wijzen van benadering. Ze zijn geschreven voor verschillende doeleinden en hebben verschillende doelgroepen.

Curriculum Vitae

- 24-09-1960 Geboren te Bad Gandersheim (Duitsland)
- 1966-1970 Lagere School te Krefeld
- 1970-1972 Gymnasium te Karlsruhe
- 1972-1979 Gymnasium Duitse School te Den Haag
- 1979-1985 Technische Universiteit Delft (Werktuigbouwkunde)
- 1985-1991 Medewerker Philips, Centrum Fabrikage Technologie, Afdeling Mechatronics
- 1991-heden Groepsleider Philips, Centrum Fabrikage Technologie, Afdeling Mechatronics, Groep Predictive Modelling

