Model-based control for high-tech mechatronic systems

Citation for published version (APA):

Document status and date:
Published: 25/11/2020

Document Version:
Accepted manuscript including changes made at the peer-review stage

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:
www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:
openaccess@tue.nl
providing details and we will investigate your claim.

Download date: 23. Sep. 2023
Model-based control for high-tech mechatronic systems

Tom Oomen
Eindhoven University of Technology

Maarten Steinbuch
Eindhoven University of Technology

CONTENTS

1.1 Introduction .......................................................... 2
  1.1.1 Motion systems .................................................. 2
  1.1.2 Industrial state of the art ...................................... 2
  1.1.3 Developments in lithography .................................... 2
  1.1.4 Developments in precision motion systems ............... 3
  1.1.5 Towards next-generation motion control: the necessity of a model-based approach .................. 4
  1.1.6 Contribution: from manual tuning to model-based synthesis ........................................... 5
  1.2 Motion systems .................................................... 6
  1.3 Feedback control design .......................................... 6
    1.3.1 System Identification - obtaining an FRF ............... 7
    1.3.2 Loop-shaping - the SISO case .............................. 8
    1.3.3 Loop-shaping - the MIMO case ............................. 9
      1.3.3.1 Interaction analysis ................................... 10
      1.3.3.2 Decoupling and independent SISO design ......... 11
      1.3.3.3 Multi-loop feedback control design with robustness for interaction .................. 11
      1.3.3.4 Sequential loop closing ................................ 12
  1.4 Model-based control design ................................... 13
    1.4.1 A standard plant approach .................................. 13
      1.4.1.1 Weighting filter selection for performance specification .................................. 14
      1.4.1.2 Obtaining a nominal model ............................ 15
      1.4.1.3 Uncertainty modeling ................................... 15
    1.4.2 Case study 1: Multivariable robust control ............ 16
    1.4.3 Case study 2: Overactuation and oversensing ......... 17
1.1 Introduction

1.1.1 Motion systems

Motion systems are mechanical systems with actuators with the primary function to position a load. The actuator can be either hydraulic, pneumatic, or electric. The freedom in trajectory planning and motion profile is often limited when compared to general robotic systems. In addition, motion systems are closely related to active vibration isolation systems, a main difference is that active vibration systems regulate a motion to zero. Examples of motion systems include linear and rotational servo drives, and also state-of-the-art planar 6 DOF (Degree-of-Freedom) motion platforms. Typically, such motion systems can be represented well by linear models, which are possibly of high order due to flexible dynamics in case of high performance and high DOF systems. Inexpensive motion systems typically have friction in the guidance. Backlash is often prevented by application of direct drive actuators. Sensor systems include encoders, capacitive sensors, and laser interferometers, with accuracies into the sub-nanometer range. For further aspects on the mechatronic design of such systems, see [38], [16].

1.1.2 Industrial state of the art

The industrial state-of-the-art control of motion systems is summarized as follows. By appropriate system design, most systems are either decoupled or can be decoupled using static input-output transformations. Hence, most motion systems and their motion software architecture use Single-Input Single-Output SISO control design method and solutions. The feedback controller is typically designed using frequency domain techniques, in particular via manual loop-shaping. A typical motion controller has a Proportional-Integral-Derivative (PID) structure, with a low-pass filter at high frequencies and one or two notch filters to compensate flexible dynamics [62] [43]. In addition to the feedback controller, a feedforward controller is often implemented with acceleration, velocity, and friction feedforward for the reference signal. The setpoint itself is designed using a setpoint generator with jerk limitation profiles [31].

One of the most accurate, fast, and expensive motion systems available today is a wafer stage, which is part of a wafer scanner used in lithographic processes. The extreme accuracy and speed requirements necessitate and justify a further development of the state of the art, which is outlined next.
1.1.3 Developments in lithography

Wafer scanners, see Fig. 1.1, are the state-of-the-art equipment for the mass production of ICs. During the production process, a photoresist is exposed on a silicon disc, called a wafer. During exposure, the image of the desired IC pattern, which is contained on the reticle, is projected through a lens on the photoresist. Subsequent chemical reactions enable etching of these patterns, which is repeated for successive layers. Typically, more than 20 layers are required for each wafer. Each wafer contains more than 200 ICs that are sequentially exposed. During this entire process, the wafer must extremely accurately track a predefined reference trajectory in six motion degrees-of-freedom (DOFs), with future requirements of tracking accuracy below 1 nm, velocity in the order of $1 \text{ m/s}$ and acceleration in the order of $10^2 \text{ m/s}^2$. These extreme speed requirements are imposed by throughput, which directly determines the market position of the machine. This highly accurate and fast motion task is performed by the wafer stage, which is one of the most accurate and expensive motion systems commercially available.

In the last decades, increasing demands with respect to computing power and memory storage have led to an ongoing dimension reduction of transistors. The minimum feature size associated with these transistors is called the critical dimension (CD) and is determined by the wavelength of light, see [34], [30]. In [30], CDs of 50 nm have been achieved using deep ultraviolet (DUV) light with a wavelength of 193 nm through many enhancements of the production process. However, a technology breakthrough is required to reduce the wavelength of light and consequently improve the achievable CD.

Extreme ultraviolet (EUV) is a key technology for next-generation lithography [65], [70]. At present, experimental prototypes with a 13.5 nm wavelength are reported in [30], [3] and production systems are already in use [1]. The introduction of EUV light sources in lithography has far-reaching consequences for all subsystems of the wafer scanner, including the wafer stage. EUV does not transmit through any known material, including air. Hence, lenses used in DUV have to be replaced by mirrors in EUV. Moreover, the entire exposure has to be performed in vacuum.

1.1.4 Developments in precision motion systems

Due to the developments in lithographic production processes, next-generation precision motion systems are expected to be lightweight for several reasons. First, vacuum operation requires these systems to operate contactless to avoid pollution due to mechanical wear or lubricants. In addition, contactless operation reduces parasitic nonlinearities such as friction and thus potentially increases reproducibility. Since contactless operation requires a compensation of gravity forces, a lightweight system is essential. Second, market viability requires a high throughput of the wafer scanner. As is argued in Section 1.1.3, this requires high accelerations in all six motion DOFs. The accelerations $a$
of the wafer stage are determined by Newton’s law $F = ma$. Here, the forces $F$ that can be delivered by the electromagnetic actuators are bounded, e.g., due to size requirements and thermal reasons. Hence, a high acceleration $a$ is enabled by a reduction of the mass $m$, again motivating a lightweight system design. Third, there have been initiatives to increase the wafer diameter from 300 mm to 450 mm to increase productivity. This requires increased dimensions of the wafer stage, which again underlines the importance of a lightweight system design.

As a result of a lightweight system design, next-generation motion systems predominantly exhibit flexible dynamical behavior at lower frequencies. This has important consequences for control design, as is investigated next.

1.1.5 Towards next-generation motion control: the necessity of a model-based approach

The increasing accuracy and performance demands lead to the manifestation of flexible dynamical behavior at lower frequencies. On the other hand, due to these increasing demands, the controller has to be effective at higher frequencies. Combining these developments leads to a situation where flexible dynamical behavior is present within the control bandwidth. This is in sharp contrast to traditional positioning systems where the flexible dynamical behavior can be considered as high-frequency parasitic dynamics, as is, e.g., the case in [71, Sec. 2.1, Assumption 1-3].

The presence of flexible dynamical behavior within the control bandwidth has significant implications for motion control design in comparison to the traditional situation:

i) next-generation motion systems are inherently multivariable, since the flexible dynamical behavior is generally not aligned with the motion DOFs;
next-generation motion systems are envisaged to be designed with many actuators and sensors to actively control flexible dynamical behavior, whereas traditionally the number of inputs and outputs equals the number of motion DOFs; and

iii) a dynamical relation exists between measured and performance variables, since the sensors generally measure at the edge of the wafer stage system, while the performance is required on the spot of exposure on the wafer itself. In contrast, the flexible dynamical behavior is often neglected in traditional motion systems, leading to an assumed static geometric relation between measured and performance variables, see Figure 1.1 for a graphical illustration.

These implications of lightweight motion systems on the control design motivate a model-based control design, since

i) a model-based design provides a systematic control design procedure for multivariable systems;

ii) a model is essential to investigate and achieve the limits of performance. Specifically, fundamental performance limitations are well-established for nominal models, see [58], and robust control provides a transparent trade-off between performance and robustness, see [60]; and

iii) a model-based design procedure enables the estimation of unmeasured performance variables from the measured variables through the use of a model.

As pointed out in [71], a model-based control design is far from standard in state-of-the-art industrial motion control, since the majority of such systems is controlled by manually-tuned single-input single-output (SISO) proportional-integral-derivative (PID) controllers in conjunction with rigid-body decoupling based on static input-output transformations. One of the main reasons is the fact that a manually tuned design achieves reasonably performance while only requiring easily available FRF data, instead of an accurate parametric model. The presence of flexible dynamical behavior in next-generation motion control necessitates and justifies the additional modelling effort required to deal with the situations sketched above.

1.1.6 Contribution: from manual tuning to model-based synthesis

The main contribution of this chapter is an overview of a systematic control design procedure for motion systems that has proven its use in industrial motion control practise. A step-by-step procedure is presented that gradually extends SISO loop-shaping to the multi-input multi-output (MIMO) situation. This step-by-step procedure consists of (i) interaction analysis, (ii) decoupling, (iii) independent SISO design, (iv) sequential SISO design, and finally, (v) norm-based MIMO design. In the norm-based MIMO design, a model-based control is pursued that addresses the future motion control challenges from Section 1.1.5. As such, the present chapter provides a unified overview of
ongoing research on motion feedback control, including [63], [50], [25], [46], and [41]. The design of the feedforward controller is an important aspect but beyond the scope of this chapter, see Section 1.5 and [43] for a results in this direction, as well as [42] for learning control.

### 1.2 Motion systems

The dynamical behavior of motion systems is typically linear and dominated by the mechanics. Indeed, the actuator and sensor dynamics are typically relevant in higher frequency regions and are therefore ignored. These mechanics are typically described in the Laplace domain as [19], [52], [37]

\[
G_m = \sum_{i=1}^{n_{RB}} \frac{c_i b_i^T}{s^2} + \sum_{i=N_{RB}+1}^{n_s} \frac{c_i b_i^T}{s^2 + 2\zeta_i \omega_i s + \omega_i^2} \tag{1.1}
\]

where \(n_{RB}\) is the number of rigid-body modes, the vectors \(c_i \in \mathbb{R}^{n_y}, b_i \in \mathbb{R}^{n_u}\) are associated with the mode shapes, and \(\zeta_i, \omega_i \in \mathbb{R}_+\). Here, \(n_s \in \mathbb{N}\) may be very large and even infinite [29]. Note that the rigid-body modes are not suspended in (1.1). In the case of suspended rigid-body modes, e.g., in case of flexures as in [16], [46], (1.1) can directly be extended. As an example, Fig. 1.2 shows a magnitude frequency response function (FRF) and the underlying modes of a system.

Traditionally, motion systems are designed to be stiff, such that the flexible behavior is well above the intended closed-loop bandwidth, implying a rigid-body behavior in the control bandwidth. Also, the number of actuators \(n_u\) and sensors \(n_y\) is chosen equal \(n_{RB}\), and are positioned such that the matrix \(\sum_{i=1}^{n_{RB}} c_i b_i^T\) is invertible. In this case, matrices \(T_u\) and \(T_y\) can be selected such that

\[
G = T_y G_m T_u = \frac{1}{s^2} I_{n_{RB}} + G_{\text{flex}}, \tag{1.2}
\]

where \(T_y\) is typically selected such that the transformed output \(y\) equals the performance variable \(z\), as is defined in Sec. 1.4.1. Importantly, the selection of these matrices \(T_u\) and \(T_y\) can be done directly on the basis of frequency response function (FRF) data, see Sec. 1.3.3.2 and, e.g., [66]. As a result of this, often decentralized controllers can be designed, i.e., diagonal PID controllers with typically a few notch filters, as is outlined in the forthcoming sections.

As is argued in Sec. 1.1.5, future motion systems are envisaged to have more actuators and sensors to improve their performance. In Sec. 1.4.3, the modal representation (1.1) is employed to actively damp and stiffen certain flexible modes.
1.3 Feedback control design

1.3.1 System Identification - obtaining an FRF

Motion control design is typically done in the frequency domain, since it allows the direct evaluation of performance and robustness in addition to intuitive tuning. Frequency domain tuning requires a model of the system. In the case where the motion system has already been realized, system identification is an inexpensive, fast, and accurate approach to model motion systems. In particular, the first step in motion systems typically involves a frequency response function (FRF) identification using noise signals, single sine, swept sine, or multi-sine excitation [51]. It is important to note that most motion systems
are unstable in open-loop, since these have rigid body dynamics leading to poles at zero in (1.1).

Indeed, due to open-loop instability, safety requirements, or nonlinear behavior, it is often required to implement a feedback controller during the identification experiments. Consider the single DOF feedback configuration depicted in Fig. 1.3. Here, $K$ can be implemented during the identification experiment, and is later re-designed to enhance system performance. For instance, $K$ can be a low bandwidth PD controller $K(s) = Ds + P$. Typically, the controller zero (ratio P/D) is taken sufficiently low. If the sign of the system at low frequencies is known, a simple procedure can be used to increase the gain, while keeping the ratio P/D constant, or first D is increased, then P, while giving the motion system a modest setpoint. In particular, a so-called jogging motion enables overcoming the friction while tuning and identifying the system. The resulting low-bandwidth controllers are typically sufficient for system identification purposes. Once these experimental controllers have been designed, an identification experiment can be performed.

Next, assume that an external excitation is applied for system identification. For the purpose of illustration, assume that $r_2 = 0$, whereas $r_1$ is the excitation signal. In the case that the measurement indeed is performed under closed-loop operating conditions, i.e., $K \neq 0$, then care must be taken when attempting to identify $G$ in Fig. 1.3. In particular, in the situation where the measured outputs are contaminated by noise and disturbances, and in addition, the external excitation $r_1$ is taken as a noise signal, as is common practice in motion control, then a direct identification of the plant, i.e., from $u$ to $y$ may lead to biased estimates. In particular, poor signal-to-noise ratios will lead to an identified inverse controller. In particular, in the case of noise excitation, it is recommended to first identify the process-sensitivity ($GS : r_1 \rightarrow y$), see again Fig. 1.3, and the sensitivity ($S : r_1 \rightarrow u$). Subsequently, divide the two FRFs to obtain the FRF of $G$. This will reduce the bias in the estimate of the plant. The same procedure can be followed for MIMO systems, provided that appropriate matrix transformations are done for every frequency measurement point. Recently, important progress has been made to enhance the FRF by appropriate input design, see [50, Appendix A], and by non-parametric pre-processing, see [69] for an application on a motion system.

In the forthcoming sections 1.3.2 and 1.3.3, the nonparametric FRF model suffices for tuning. In Sec 1.4, a parametric model is required. This parametric model is estimated based on the obtained FRF data.

1.3.2 Loop-shaping - the SISO case

The feedback controller can be directly designed based on the obtained FRF. First, manual loop-shaping for SISO systems is investigated. As an example consider the measured FRF of a wafer stage that is presented in Section 1.4.2, see Figure 1.6.

The idea in loop-shaping is to select $K$ that directly affects the loop-gain
$L = GK$. Indeed, $L$ can be directly manipulated by the choice of $K$ using Bode and Nyquist diagrams, see, e.g., [4, Sec. 11.4], which is essentially due to the Laplace transform that replaces the time domain convolution by a simple multiplication. In turn, $L$ directly connects to the closed-loop transfer functions in certain frequency ranges, including $S = \frac{1}{1+GK}$ and $T = \frac{GK}{1+GK}$, see also Fig. 1.3. Loop-shaping typically consists of the following steps, which are typically re-tuned in an iterative fashion.

1. Pick a cross-over frequency $f_{bw}$ [Hz], which is the frequency where $|L(2\pi f_{bw})| = 1$. Typically, $|S| < 1$ below $f_{bw}$, while $|T| < 1$ beyond $f_{bw}$. Furthermore, $f_{bw}$ is typically selected in the region where the rigid-body modes dominate, i.e., $f_{bw} < \frac{1}{2\pi}$, where $\omega_i, i = n_{rh} + 1, \ldots, n_s$ is defined in (1.1).

2. Implement a lead filter, typically specified as $K_{lead} = p_{lead} = \frac{\frac{1}{3} f_{bw} + 1}{\frac{2\pi}{f_{bw}} + 1}$, such that sufficient phase lead is generated around $f_{bw}$. In particular, the zero is placed at $\frac{1}{3} f_{bw}$, while the pole is placed at $3 f_{bw}$. Next, $p_{lead}$ is adjusted such that $|GK_{lead}(2\pi f_{bw})| = 1$. This lead filter generates a phase margin at $f_{bw}$. Indeed, the phase corresponding to the rigid-body dynamics in (1.1) is typically 180 degrees, which leads to a phase margin of 0 degrees.

3. Check stability using Nyquist plot of $GK$ and include possible notch filters in the case where the flexible modes endanger robust stability.

4. Include integral action through $K = K_{int} K_{lead}$, with $K_{int} = \frac{s + 2\pi f_{bw}}{s}$. Since $GK \gg 0$ at low frequencies, $S \approx \frac{1}{GK}$, the integrator pole at $s = 0$ improves low frequency disturbance attenuation, while not affecting phase margin due to the zero at $\frac{1}{3} f_{bw}$.

The loop-shaping procedure is highly systematic and fast: it only requires a nonparametric FRF model, which is fast, inexpensive, and accurate, and the above procedure is very systematic and intuitive to apply. However, in case the system is multivariable, caution should be taken to deal with interaction.

### 1.3.3 Loop-shaping - the MIMO case

Loop-shaping cannot be directly applied to MIMO systems due to interaction. Essentially, the interdependence between the decentralized PID controllers complicates the tuning of the overall multivariable controller that at least has to be stabilizing. In this section, the manual loopshaping procedure is gradually extended to the multivariable case. These results are essentially all based on the generalized Nyquist criterion.

In MIMO systems, closed-loop stability is determined by the closed-loop characteristic polynomial $\det(I + L(s))$. By graphically analyzing $\det(I + L(j\omega))$ similar to the scalar Nyquist plot, closed-loop stability can be analyzed, see [59, Theorem 4.9] for details. This plot can be generated directly using the identified MIMO FRF of Sec. 1.3.1. However, in case the closed-loop systems is unstable, it is not immediate which element of the MIMO controller $K$ to re-tune.
Note that a related test can be obtained in terms of the loci of \( L(j\omega) \), i.e., the eigenvalues of \( L \) as a function of frequency. In case both \( K \) and \( G \) are open-loop stable, then a sufficient condition for closed-loop stability is where all loci do not encircle the point \((-1,0)\). However, this does not resolve the design issue: the shaping of these eigenvalue loci is not straightforward if the plant has interaction. Furthermore, typical margins that are used in SISO systems, such as phase margins are less useful. In particular, the phase margin based on characteristic loci implies a phase change at the same time in all loops simultaneously.

In the special case where the open loop transfer function matrix is diagonal, \( L(s) = \text{diag}\{l_i(s)\} \), i.e. the open loop is decoupled, then
\[
\det(I + L(s)) = \prod_{i=1}^{n} \det(1 + l_i(s)). \tag{1.3}
\]
As a result, the characteristic loci of the open loop transfer function matrix directly coincide with the Nyquist plots of the scalar loop gains \( l_i(j\omega) \). The MIMO feedback control design complexity then reduces to that of a number of SISO feedback control design. Many classical MIMO control design methods aim at decoupling the open loop function at some location in the feedback loop, e.g. at the plant input or plant output.

Since SISO loop-shaping is systematic and very fast design procedure for motion systems, it is a preferred procedure even for MIMO systems. In practice, it is often attempted to decouple the system such that (1.3) holds. If this is not possible, it may be tempting to directly go to a norm-based optimal design, however, this requires a parametric model. Such parameteric models, as is outlined in Sec. 1.4 are expensive and user-intensive to obtain. Therefore, several ideas from multivariable and robust control are exploited to extend loop-shaping towards MIMO systems. This leads to a design procedure for MIMO motion systems, consisting of the following steps.

1. interaction analysis,
2. decoupling and possible independent SISO design,
3. multi-loop feedback control design with robustness for interaction,
4. sequential feedback control design,
5. norm-based control design.

All except for the last step can be performed with a non-parametric model of the plant, i.e., an identified FRF. The norm-based control design requires a parametric model of the plant. Nonetheless, the last step may be essential to address the complex motion control problems envisaged in future systems, see Sec. 1.1.5.
1.3.3.1 Interaction analysis

The first step in multivariable motion feedback design is interaction analysis. Especially two-way interaction is essential, since one-way interaction does not affect closed-loop stability.

Several approaches have been developed for interaction analysis. A well-known, easy to compute, and useful approach is the relative gain array. The frequency-dependent relative gain array (RGA) \([60, 11]\) is given by

\[
\text{RGA}(G(f)) = G(f) \times (G(f)^{-1})^T,
\]

where \(\times\) denotes element-wise multiplication. Note that the RGA can be directly computed using FRF data. In addition, the RGA is independent of the feedback controller and invariant under scaling. The rows and columns of the RGA sum to 1 for all frequencies \(f\) (Hz). If \((\text{RGA})(f) = I, \ \forall \omega,\) then perfect decoupling is achieved.

1.3.3.2 Decoupling and independent SISO design

Static decoupling may be considered if the interaction is too severe to allow for multi-loop SISO design. For motion systems, such transformations, see also (1.2), can be obtained using kinematic models. Herein, combinations of the actuators are defined so that actuator variables act in independent directions at the center of gravity. Similarly, combinations of the sensors are defined so that each translation and rotation of the center of gravity can be measured independently. Such decouplings can be further refined using FRF data, see [66] for details.

In certain situations, it may be desirable to decouple the plant at other frequencies, see, e.g., [8], or to use a dynamic decoupling. In any situation, the effect of the decoupling transformations should be analysed using the interaction measures derived earlier. If the system is sufficiently decoupled using static transformations, then multi-loop SISO controllers can be designed.

1.3.3.3 Multi-loop feedback control design with robustness for interaction

In case where the interaction after decoupling is too large to successfully design multi-loop SISO controllers, robustness for interaction may be enforced. This approach employs concepts from robust control theory [72]. The objective is to design SISO controllers that are robust for interaction terms.

To this end, note that [23]

\[
\det(I + GK) = \det(I + E_T T_d) \det(I + G_d K),
\]

with, \(T_d = G_d K (I + G_d K)^{-1}\). Let \(G_d(s)\) be the diagonal terms of \(G\). Then, the non-diagonal terms of the plant \(G_n d(s) = G(s) - G_d(s)\) can be considered in \(E_T(s)\). As a result, MIMO closed-loop stability can be decomposed into
stability of \( N \) non-interacting loops associated with \( \det(I + G_d(s)K(s)) \) and stability associated with \( \det(I + E_T(s)T_d(s)) \).

In \( \det(I + E_T(s)T_d(s)) \), \( T_d \) is the complementary sensitivity function of the \( N \) decoupled loops. If \( G(s) \) is stable and \( T_d(s) \) is stable, then the small gain theorem [59] implies that \( \det(I + E_T T_d) \) does not encircle the origin if,

\[
\rho(E_T(j\omega)T_d(j\omega)) < 1, \forall \omega, \tag{1.6}
\]

where \( \rho \) is the spectral radius. Next, (1.6) holds if

\[
\mu_{T_d}(E_T(j\omega)T_d(j\omega)) < 1, \forall \omega \tag{1.7}
\]

which in turn implies that

\[
\sigma(T_d(j\omega)) < \mu_{T_d}(E_T(j\omega))^{-1}, \forall \omega \tag{1.8}
\]

where \( \mu_{T_d} \) is the structured singular value [72] with respect to the diagonal (decoupled) structure of \( T_d \). Since \( \sigma(T_d(j\omega)) = \max_i |T_{d,ii}(j\omega)| \), condition (1.8) can be used as an additional bound on \( T_{d,ii} \) when loop-shaping the multi-loop SISO controllers. In other words, once the structured singular value in (1.8) is computed, the SISO approach in Sec. 1.3.2 can directly be followed, where an additional bound is enforced on all SISO complementary sensitivity functions \( T_{d,ii} \) to provide robustness for interaction terms.

Clearly, the test (1.8) provides robustness for any choice of stabilizing controller in the other loops. Intuitively, this implies that this may be conservative. In the next section, the particular choice of controller in a previous loop is explicitly taken into account, leading to a procedure for designing decentralized controllers for interaction.

### 1.3.3.4 Sequential loop closing

In the situation where the approach in Sec. 1.3.3.3, which provides robustness for interaction, leads to conservative designs, then the interaction terms and specific controller designs may be explicitly taken into account in subsequent design steps. This procedure is called sequential loop closing (SLC). Here, earlier designed controllers are explicitly taken into account in the next loops. The idea is to start with an FRF of the open-loop system, which is assumed to be open-loop stable. Then, the first loop is closed, and a new FRF is computed of the resulting system, which includes effects of the designed controller. Then, one more loop is closed, and the resulting FRF is computed again. In particular, each single input single output controller \( k_i \), from \( K = \text{diag}\{k_i\} \), \( i = \{1, ..., n\} \), is designed using the property [35]

\[
\det(I + GK) = \prod_{i=1}^{n} \det(1 + g^i k_i), \tag{1.9}
\]

where for each \( i^{th} \) design step, the equivalent plant \( g^i \) is defined as the lower
fractional transformation:

\[ g^i = F_i(G, -K^i) \]  

where \( K^i = \text{diag}\{k_j\}, j = \{1, ..., n\}, j \neq i. \)

It can be directly proved that the closed-loop system is stable if every \( k_i \) is designed to be stabilizing. However, the robustness margins, that are typically enforced for each \( k_i \), may change due to the closing of subsequent loops. As a consequence, the overall MIMO design may have very poor robustness margins [17].

To alleviate these drawbacks, the ordering of the design steps may be changed, since this may significantly influence the achieved performance. There is no general approach to determine the best sequence for design. This may lead to many design iterations, especially for large MIMO systems. Note that the loops should always be opened in the same ordering as they are closed, otherwise closed-loop stability cannot be guaranteed.

The main advantage of the sequential loop closing approach is that it allows for SISO loop-shaping while explicitly addressing interaction, while only requiring a non-parametric plant model. If the sequential loop-closing approach still does not lead to satisfactory performance, then a full multivariable control design approach may be pursued, as is investigated next. Such an approach may become essential for envisaged future systems as described in Sec. 1.1.5, which exhibit a drastically increased complexity.

### 1.4 Model-based control design

A full model-based control design is a suitable approach in the case where the manual tuning approaches that are presented in the previous section do not lead to satisfactory results. In this section, a general framework for model-based motion control is outlined. The presented framework relies on parametric models of the system and optionally a detailed description of the involved uncertainty. Due to the need for parametric models and uncertainty models, the proposed model-based control framework requires a significantly larger investment of time and effort in obtaining such models. Relevant situations where the additional investment may be justified and may be necessitated include, see also Sec. 1.1.5:

- multivariable systems with large interaction and/or uncertainty;
- overactuation and oversensing, i.e., the use of additional actuators and sensors to actively control flexible dynamics; and
- inferential control, i.e., the situation where not all performance variables are measured variables for feedback control.
1.4.1 A standard plant approach

A standard plant approach is pursued to deal with the relevant situations mentioned in Sec. 1.4. The considered standard plant is depicted in Fig. 1.4. Here, $P$ denotes the standard plant and $C$ is the to-be-designed controller. In addition, the vector $w$ contains exogenous input variables, including disturbances, sensor noise, and setpoints, whereas the vector $z$ contains the regulated variables that are defined such that they are ideally zero, e.g., servo errors. Furthermore, the vector $y$ contains the measured variables that are available to the feedback controller, whereas the vector $u$ contains the manipulated variables.

The standard plant is general in the sense that enables the formulation of general control problems, including an arbitrary number of regulated variables $u$ and measured variables $y$ for overactuation and oversensing, and a distinction between measured variables $y$ and performance variables $z$ as is required in the inferential control situation.

Once the motion control problem is posed in the standard plant framework, a large amount of literature, including [60] and [72], and software toolboxes, including [5], are available to synthesize the optimal controller $C$. Posing the control problem properly into the standard plant involves at least three aspects that are investigated next: 1. weighting filter selection, 2. obtaining a parametric model of the motion system, and 3. uncertainty modeling.

1.4.1.1 Weighting filter selection for performance specification

To properly pose the control problem, weighting filters have to be selected. A criterion of the form

$$J(P, K) = \|F_l(P, K)\|$$

(1.11)

is adopted, where the goal is to compute

$$K^{\text{opt}} = \min_K J(P, K).$$

(1.12)

Here, a specific norm $\|\cdot\|$ has to be selected, e.g., $\mathcal{H}_2$ or $\mathcal{H}_\infty$. For motion control problems, the $\mathcal{H}_\infty$ norm has proven to be particularly useful, since it enables
the specification of performance weights through loop-shaping ideas, see Section 1.3.2 and Section 1.3.3, and it enables the design of robust controllers by explicitly addressing model uncertainty. The $H_\infty$ norm for a stable system $H$ is given by
\begin{equation}
\|H\|_\infty = \sup_\omega \sigma(H(j\omega)).
\end{equation}

The next step in specifying the performance goals is to select the internal structure of the standard plant. For instance, in the case where $y = z$ a common choice is
\begin{equation}
F_l(P,K) = W \begin{bmatrix} G \\ I \end{bmatrix} (I + KG)^{-1} \begin{bmatrix} K \\ I \end{bmatrix} V.
\end{equation}
In (1.14), $G$ denotes the open-loop system and $W$ and $V$ are weighting filters. These weighting filters enable the specification of various control goals and have to be user-specified. When using the $H_\infty$ norm, these can be specified by means of desired loop-shapes of closed-loop transfer function matrices [60] or open-loop gains [67]. Suitable weighting filters for typical motion control problems that are in line with the guidelines in Section 1.3.2 and Section 1.3.3 have been developed in, e.g., [56], [62], whereas extensions to multivariable systems are presented in [71] and [8].

1.4.1.2 Obtaining a nominal model

Besides the specification of the control goal through weighting functions, the optimization in (1.12) requires knowledge of the true system denoted by $G_0$. In a model-based control design, the knowledge of the true system is reflected by means of a model $\hat{G}$. System identification is an accurate, fast, and inexpensive approach to obtain the required model for motion systems. To obtain an accurate model, the model should be tailored towards the control goal [21], [57]. These results are further developed in [12] and [47] towards the identification of low-order models for control design. The main ingredient for these algorithms is the control criterion (1.11) and closed-loop FRFs, as is explained in Section 1.3.1.

1.4.1.3 Uncertainty modeling

The model $\hat{G}$ obtained in through the procedure in the previous section is not an exact description of the true system $G_0$, since i) motion systems generally contain many resonance modes [29] of which a limited number is included in the model; ii) parasitic nonlinearities are present, e.g., nonlinear damping [61]; and iii) identification experiments are based on finite time disturbed observations of the true system.

Robust control design [72], [60] explicitly addresses these model errors by considering a model set by extending the nominal model $\hat{G}$ with a description of its uncertainty. This model set has to be chosen judiciously, since the resulting robust controller is typically designed to optimize the worst-case
performance over all candidate models in this set. This has led to the development of model uncertainty structures for robust control, where the traditional uncertainty structures, e.g., [72, Table 9.1], have been further developed towards coprime factor perturbations [36], [32], and dual-Youla structures [2], [12], [14], see [45] for recent developments.

1.4.2 Case study 1: Multivariable robust control

To illustrate the concepts of Sec. 1.4.1, a model-based controller is designed for a traditional motion control problem, i.e., non-inferential ($y = z$) and non-overactuated/non-oversensed ($\dim y = \dim u$, which in turn equals the number of motion degrees of freedom). In particular, the $x$ and $y$ translational degrees of freedom of the wafer stage in Fig. 1.5 are considered. This wafer stage is the main high-precision motion system in a lithographic machine [3].

The following sequence of steps is applied.

1. An FRF is identified following the procedure in Sec. 1.3.1.

2. Weighting filters $W$ and $V$ are specified as outlined in Sec. 1.4.1.1. In view of the procedure in Section 1.3.2, a cross-over frequency of 90 Hz is set as target for both the $x$ and $y$ translational degrees-of-freedom.

3. A nominal model $\hat{G}$ is identified using the procedure outlined in Sec. 1.4.1.2. The resulting model is depicted in Figure 1.6. The model is control-relevant and of low order. Indeed, the model is accurate around the desired cross-over region where the rigid body modes are accurately modeled. In addition, the first two resonance phenomena around 200 Hz are captured into the model, since these are known to endanger stability. In addition, the model is of order 8: 4 states correspond to two rigid body modes, whereas the other 4 states correspond to the inherently multivariable flexible modes around 200 Hz.

4. The uncertainty of the nominal model is taken into account, see Sec. 1.4.1.3. Here, the emphasis is on selecting the uncertainty such that the true system behavior $G_o$ is encompassed but that the uncertainty is selected as parsimonious as possible. The resulting model set is depicted in Fig. 1.6. Clearly, the model set is extremely tight in the cross-over region, which is well-known to be important for subsequent control design [6].

5. Next, a robust controller $K$ is synthesized using commercially available optimization algorithms [5]. Note that the resulting controller $K$ is multivariable and obtained in one shot, which is in sharp contrast to the approach in Sec. 1.3.3.4. As a result, optimality is guaranteed for the multivariable system.

6. The resulting controller is implemented on the experimental system, see Figure 1.7 for the resulting power spectra. When compared to a standard
manually tuned PID controller, the optimal robust controller automatically generates notch filters and leads to a factor four in variance reduction. Further details of the approach and experimental results are presented in [50].

1.4.3 Case study 2: Overactuation and oversensing

Besides the fact that the approach in Sec. 1.4.1 can deal with the traditional motion control design problem as shown in Sec. 1.4.2 and [71], it can also systematically address control problems that are not immediate to solve by manual design. An important case is the situation where additional actuators and sensors are used to enhance control performance. Indeed, when the sensors and actuators are equal to the number of motion degrees of freedom, the achievable performance is limited, see [8] for a systematic $H_\infty$-optimal approach to achieve optimal performance for this situation and [28] for a manual loop-shaping based approach. The main rationale in overactuation and oversensing is that additional actuators and sensors, i.e., strictly more than motion DOFs, can significantly enhance performance.

The basic idea is to extend the setup of Figure 1.4 towards Figure 1.8, where additional inputs $u_{\text{ext}}$ and outputs $y_{\text{ext}}$ are introduced. In particular,
FIGURE 1.6
Bode magnitude diagram of FRF (blue dots), nominal model $\hat{G}$ (solid blue) and uncertain model set for robust control (cyan).
Consequently, the resulting system identification and robust unstructured uncertainty without introducing conservatism. The key step that enables the results in this paper involves (dashed red).

FIGURE 1.7
Cumulative power spectrum of the measured error in x-direction (left) and y-direction (right): manually tuned controller for reference (solid blue) and optimal robust controller (dashed red).

FIGURE 1.8
Extended standard plant configuration with additional control inputs and outputs.
Model-based control for high-tech mechatronic systems

Consider the situation in Figure 1.8, then it can be immediately verified that

\[
\begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 0
\end{bmatrix}
\begin{bmatrix}
P_{\text{ext}} \\
I \\
0
\end{bmatrix}
= P.
\]

(1.15)

with corresponding performance \( J_{\text{ext}} \), which is achieved by a controller \( K_{\text{ext}} \) in the extended configuration, is given by

\[
J_{\text{ext}}(K_{\text{ext}}) := \| \mathcal{F}(P_{\text{ext}}, K_{\text{ext}}) \|.
\]

(1.16)

The rationale behind the setup in Figure 1.8 is that the bound

\[
\min_{K_{\text{ext}}} J_{\text{ext}}(K_{\text{ext}}) \leq J(K^*)
\]

(1.17)

holds by definition, which can be directly seen by setting \( C_{\text{ext}} = \begin{bmatrix} K^* & 0 \\ 0 & 0 \end{bmatrix} \) it holds that \( J_{\text{ext}}(K_{\text{ext}}) = J(K^*) \). By performing the procedure of Sec. 1.4, it can be achieved that

\[
\min_{K_{\text{ext}}} J_{\text{ext}}(K_{\text{ext}}) \ll J(K^*)
\]

(1.18)

if appropriate design and theoretical considerations are followed as suggested in [25].

A wafer stage case study is presented to show potential performance improvement of overactuation and oversensing. Herein, control design in the vertical direction is considered, i.e., the translational direction z and rotations \( R_x \) and \( R_y \) as depicted in Figure 1.9(a). Also, 4 actuators are placed as indicated in Fig. 1.9(a). Specifically, \( a_1, a_2, \) and \( a_3 \) are placed below the corners of the stage, whereas \( a_4 \) is positioned at the middle of the line between the center and \( a_2 \). Moreover, corner sensors \( s_1, s_2, \) and \( s_3 \) are used as indicated in Fig. 1.9(a). Finally, a piezo sensor \( s_4 \), which measures strain of the wafer stage, is available at the middle of the line between the center and \( s_2 \). Since 4 actuators and sensors are available to control the 3 rigid-body DOFs, there is freedom left to actively control flexible dynamical behavior of the wafer stage.

An extended version of the procedure in Section 1.4.2 is followed. In particular, the system is rigid-body decoupled except for an additional input-output pair that is available for control of relevant flexible dynamical behavior. The first flexible mode of the system is a torsion bending of the stage, depicted in Fig. 1.9(b). This mode can effectively be controlled using the fourth input-output direction of \( P_{\text{ext}} \), which is indicated by \( P_{\text{flex}} \) in Fig. 1.10. On the one hand, the fourth output of \( P_{\text{ext}} \) is the piezo sensor \( s_4 \), which measures strain. As a consequence, no rigid-body displacements are observed at this output, while the structural deformations associated with the flexible dynamical behavior of the system are measured indeed. On the other hand, the fourth input of \( P_{\text{ext}} \) does not excite rigid-body behavior of the system.
The main idea is that the fourth loop can be closed to actively compensate the torsion mode in Fig. 1.9(b). To illustrate this, the equivalent plant as defined in (1.10) is depicted in Fig. 1.11. The main result is an increase of the torsion mode frequency from 143 Hz to 193 Hz as well as an increased damping of this loop. This enables enhanced performance for the remaining equivalent plant, since this torsion mode was performance limiting due to non-collocated dynamical behavior. In particular, active control of the torsion mode implies that a higher cross-over frequency can be achieved in the original motion degrees of freedom.

To reveal the potential performance enhancement with additional actuators, a controller for the conventional situation in Figure 1.4 to the extended control configuration in Figure 1.8, see also Fig. 1.11. The resulting controllers are implemented on the wafer stage system and stand-still errors are measured, see Fig. 1.12. Clearly, the performance has been significantly increased by active control of the torsion mode. Further details of this approach are reported in [26], see also [55] for a related feedforward control approach.

1.4.4 Case study 3: inferential control
As is argued in Section 1.1.5, increasing performance requirements lead to a situation where the measured variables $y$ are not a good representation of the performance variables $z$. In motion systems, mechanical deformations may be present between the location where performance is desired and where the measurement takes place. This is shown in Fig. 1.1. In this section, a third case study is presented to show potentially poor controller designs when this is not properly addressed. In addition, by means of a model-based approach, the performance variable $z$ can be (possibly implicitly) inferred from the measured variables $y$.

The high-precision prototype motion system in Fig. 1.13 has been developed for evaluating control strategies in next-generation motion systems that exhibit dominant flexible behavior. Four out of six motion DOFs of the movable steel beam have been fixed by means of leaf springs. Thus, the system can
FIGURE 1.10
Identified FRF $G_{o,ext}$ and control-relevant 8th order parametric model $\hat{G}_{ext}$.

FIGURE 1.11
Equivalent plant $G_{o,eq}$ (dotted) and parametric fit $\hat{P}_{eq}$ (solid) under active control of the torsion loop. The initial model of motion DOFs $[z_R, R_x, R_y]$ Fig. 1.10 is shown for comparison (dashed).
FIGURE 1.12
Cumulative power spectral density (CPSD) for the conventional controller of Figure 1.4 (blue) and for the extended control configuration controller of Figure 1.8.

FIGURE 1.13
Photograph of the experimental flexible beam setup with three actuators and three sensors.

FIGURE 1.14
Schematic top view illustration of flexible beam setup.
move $x$ and $\varphi$ direction indicated in Fig. 1.14. The inputs consist of 3 current-driven voice-coil actuators, whereas the outputs are 3 contactless fiberoptic sensors with a resolution of approximately $1 \, \mu\text{m}$. The following selection is made to show possible hazards in the inferential control situation and a solution. The goal is to control the translation of the center of the beam, i.e., at sensor location $s_2$, hence

$$z_p = s_2,$$

(1.19)

where $s_2$ is unavailable for the feedback controller and $z_p$ will be further explained below. Regarding the measured variable, the center of the beam is determined by averaging the outer sensors $s_1$ and $s_3$, i.e.,

$$y_p = \left[ \frac{1}{2} \quad \frac{1}{2} \right] \begin{bmatrix} s_1 \\ s_3 \end{bmatrix},$$

(1.20)

which in fact corresponds to a sensor transformation based on static geometric relations as is indicated in Fig. 1.14. Consequently, a discrepancy between the measured variable $y_p$ and performance variable $z_p$ may exist due to internal deformations of the beam. Only the outer actuators are used for control, i.e.,

$$a_1 = a_3 = u_p, \quad a_2 = 0.$$  

(1.21)

Comparing (1.20) and (1.21) reveals that $u_p$ and $y_p$ are collocated. The resulting system is given by

$$\begin{bmatrix} z_p \\ y_p \end{bmatrix} = \begin{bmatrix} P_x \\ P_y \end{bmatrix} u_p = P u_p,$$

(1.22)

where $z_p$ denotes the point of interest of the system, $y_p$ is the measured variable available for feedback control, and $u_p$ is the manipulated plant input. Note that in the standard plant formulation of Fig. 1.4, $z_p$ will be part of $z$, in addition to, e.g., the input signal to avoid excessive control inputs, see [46] for details. The resulting control problem above has closely resembles the example in Fig. 1.1 and standard geometric decoupling techniques, e.g., [71].

Next, the procedure in Section 1.4.2 is followed again, i.e., it is for the moment assumed that $y_p = z_p$. The resulting controller is implemented on the beam system, where the measured response $y_p$ is shown in Fig. 1.15 (bottom). The response is $y_p$ as expected from simulations and sufficiently fast to reach the setpoint. However, when inspecting the measurement $z_p$, which is not available to the feedback controller, but measured for performance validation, it is clear that system exhibits large internal deformations. Hence, the true performance of the system is poor if one neglects the fact that a distinction exists between $z_p$ and $y_p$. In addition, this poor performance cannot be directly observed from the feedback loop signals.

In [46], an inferential motion control design framework is presented. This framework exploits measurements $y_p$ and a model of $P_z$ to infer the performance variables, i.e., it does not require a real-time measurement of the
1.5 Conclusion

In this chapter, a step-by-step model-based control design approach is presented that is applied to several high-tech industrial case studies. The main rationale of the design approach is to increase the complexity of the control design procedure and the associated modeling approach only if justified and necessitated by the performance and accuracy requirements. The step-by-step procedure consists of the following steps: (i) interaction analysis, (ii) decoupling, (iii) independent SISO design, (iv) sequential SISO design, and (v) norm-based MIMO design. The required model ranges from inexpensive FRF data in step (i)-(iv) towards a parametric model in step (v).

performance variables $z_p$. This framework is applied to design an inferential motion controller for the beam setup in Figure 1.13 and Figure 1.14. The results are also depicted in Figure 1.15. Clearly, the performance in terms of $z_p$ significantly increases. For further details, the reader is referred to [46].

FIGURE 1.15
Experimental step responses: inferential motion controller (solid blue), traditional robust control design (dashed red).
Three recent case studies are presented where the performance and accuracy requirements lead to a complex control design problem that necessitates a model-based optimal design. These include (i) a robust multivariable control design for a traditional yet high-performance wafer stage control problem, (ii) a case study on a prototype wafer stage to exploit overactuation and oversensing, and (iii) a case study on a prototype motion system to address an inferential control problem where the performance variables are not available for real-time feedback control.

Ongoing research mainly focuses modeling. Regarding steps (i) - (v), non-parametric identification is being enhanced to improve the FRF quality. This includes the use of multi-sine excitation signals, see [50, Appendix A] for initial results, and the use of preprocessing using LRM/LPM, see [69], [33], [15] for initial results that are presently being applied to multivariable FRF estimation.

Regarding the norm-based MIMO control design step (v), the required modeling effort is a major obstruction for industrial application. Several aspects are important. First, regarding modeling for robust control, present research focuses on new approaches to identify ‘good’ model sets for robust control. For mechatronic systems, important aspects include the identification of a control-relevant nominal model, e.g., [57], the structure of model uncertainty, see, e.g., [14] and [32] for general results, and [45] for recent developments, and the size of model uncertainty, approaches that are particularly suitable for mechatronic systems include, e.g., [20] for local parametric models, [49] for a data-driven approach, and [44] for a validation-based approach.

Second, the identification of accurate models of complex high-tech systems is numerically challenging. To enable the accurate and fast identification of nominal models, new approaches that enable a numerically reliable implementation are being developed, see, e.g., [27], [24], see also [68] for a recent benchmark comparison.

Third, motion systems lead to position-dependent behavior due to the fact that the system inherently makes movements. At present, methods are being developed to quantify the position dependence through control-relevance, see [39, Section 5.7] for early results, as well as explicitly modeling the position dependence by exploiting the physical model structure (1.1) in an LPV framework, see also [64] and [22] for general LPV modeling and motion control. Furthermore, identification of more general non-linear behavior is investigated in, e.g., [54].

Fourth, disturbances have a crucial role in disturbance attenuation, see [9] for initial research in this direction. Ongoing research focuses on control-relevant disturbance modeling in the framework of [45].

Fifth, sampled-data identification and control is being investigated. All the results presented in this chapter have tacitly been developed in the discrete time domain. This neglects the important issue of intersample behavior, see [13] for general results and [48] for a motion control framework. Such sampled-data frameworks necessitate the modeling for sampled-data control,
Acknowledgement

essentially requiring continuous time models [18]. In [40], a multirate framework is presented that addresses this to best extent for periodic sampling. Note that periodic sampling will always be restricted due to the Nyquist-Shannon sampling theorem, yet at a higher frequency bound. This bound is for motion systems typically fairly high. If this is not sufficient, the approach can be further enhanced by measuring irregularly sampled data, e.g., [18], which is straightforward to perform in certain motion systems.

Finally, the present chapter focussed on feedback control design. For high performance motion control, feedforward is equally important. For traditional feedforward tuning, the reader is referred to [63, Section 27.3], [43]. Ongoing research includes iterative learning control and automated feedforward control tuning, see [42], both based on a tailored iterative learning control approach, see [10], and based on instrumental variable system identification, see [7]. Related approaches are presented in, e.g., [53], where a frequency domain approach is presented for nonlinear systems.

1.6 Acknowledgement

The authors would like to thank our very many industrial and academic collaborators, as well as former students, who have contributed to the overall framework and experimental results presented in this chapter. This work is part of the research programme VIDI with project number 15698, which is (partly) financed by the Netherlands Organisation for Scientific Research (NWO).
Bibliography


