





Chapter 3

Our Specific Approach on Mastering Uncertainty



Peter F. Pelz , Robert Feldmann, Christopher M. Gehb, Peter Groche , Florian Hoppe, Maximilian Knoll, Jonathan Lenz, Tobias Melz, Marc E. Pfetsch , Manuel Rexer, and Maximilian Schaeffner 

Abstract This chapter serves as an introduction to the main topic of this book, namely to master uncertainty in technical systems. First, the difference of our approach to previous ones is highlighted. We then discuss process chains as an important type of technical systems, in which uncertainty propagates along the chain. Five different approaches to master uncertainty in process chains are presented: uncertainty identification, uncertainty propagation, robust optimisation, sensitivity analysis and model adaption. The influence of the process on uncertainty and methods depends on whether it is dynamic/time-varying and/or active. This brings us to the main strategies for mastering uncertainty: robustness, flexibility and resilience. Finally, three different concrete technical systems that are used to demonstrate our methods are presented.

How can we ensure product safety, realise lightweight structures, enable sustainable systems or control production quality? All these questions lead to the issue of how to master uncertainty, cf. Chap. 1. The answer to this core issue has many different facets, and there is a need to structure the topic and demonstrate solution methods.

This chapter is dedicated to structuring the facets, but also to introducing our specific approach to mastering uncertainty. For this purpose, we introduce process chains and temporal characteristics of processes. Often, active processes seem to be a smart way to cope with uncertainty. We compare the cost benefit associated with the selection of active processes.

The original version of this chapter was revised with the missed out corrections. An erratum to this chapter can be found at https://doi.org/0.1007/978-3-030-78354-9_8

P. F. Pelz (✉) · R. Feldmann · Christopher M. Gehb · P. Groche · F. Hoppe · M. Knoll · J. Lenz · T. Melz · M. Rexer · M. Schaeffner
Department of Mechanical Engineering, TU Darmstadt, Darmstadt, Germany
e-mail: peter.pelz@fst.tu-darmstadt.de

T. Melz
Fraunhofer Institute for Structural Durability and System Reliability LBF, Darmstadt, Germany

Marc E. Pfetsch
Department of Mathematics, TU Darmstadt, Darmstadt, Germany

© The Author(s) 2021, corrected publication 2021
P. Pelz et al. (eds.), *Mastering Uncertainty in Mechanical Engineering*,
Springer Tracts in Mechanical Engineering,
https://doi.org/10.1007/978-3-030-78354-9_3

The three strategies, (i) robustness, (ii) flexibility, (iii) resilience, are of particular interest when it comes to mastering uncertainty. In this chapter we introduce these and distinguish between the three strategies. Finally, to demonstrate and validate our specific approach, we define three demonstrator systems in the last section of this chapter.

3.1 Beyond Existing Approaches

Peter F. Pelz

Uncertainty quantification is a large research field today. It is the basis for machine learning and inference in the context of data-driven, i.e. black box models, cf. Chap. 1. Uncertainty quantification inevitably deals with model uncertainty and data uncertainty. Grey box models become more and more important in mechanical engineering, where prior knowledge is available.

Our research from the perspective of mechanical engineering has three aspects that go beyond the field of uncertainty quantification:

- (i) Firstly, we focus on physical systems that are designed, manufactured and used. The materiality of the systems and the physicality of their use inevitably lead to new aspects in the mastering of uncertainty. One example is the mastering of uncertainty across all product life phases and the concept of closed loops between phases, Fig. 3.1. Another example is the consideration of the functional safety of physical systems.
- (ii) Secondly, the decomposition into sub-functions and the assignment of real processes or components to the sub-functions, typical for mechanical engineering [36], is a guiding principle for us, which is reflected in the uncertainty classification given in Chap. 2. It leads us to the concept of structural uncertainty.

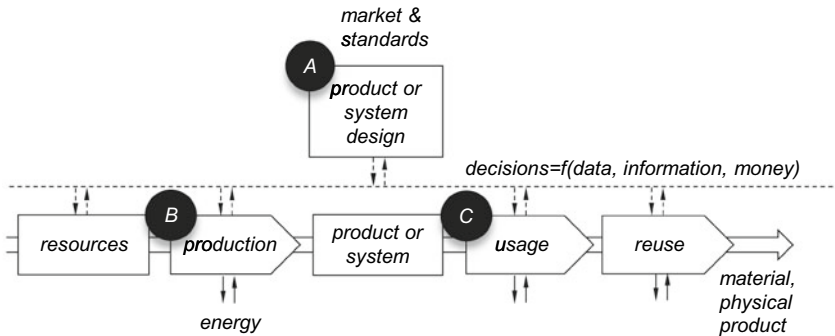


Fig. 3.1 A product or system design, B production, C usage; all phases are interconnected by the flow of material and physical products as well as data, information and money

But the systematic engineering design view also leads us to the conflicting objectives minimum effort, maximum availability and maximum acceptability. This view results in the broad motivation of mastering uncertainty. It is a basic approach to gain product safety, lightweight structures, sustainable systems and closed-loop production quality control, Chap. 1.

- (iii) Thirdly, and most importantly in our view, we focus on mastering uncertainty; quantifying it may be a prerequisite. As a result, we put strategies for uncertainty management on the top level. Some of these strategies require new active or semi-active technologies that enable adaptation to changes in production or usage. Here, too, we extend the current state of science and technology by our contribution.

View from different disciplines

In the twelve years of research that form the foundation of this book, there has been continuous inspiration between the various disciplines. Thus, the framework structure outlined in Fig. 1.12 can be viewed equally from the perspectives of mechanical engineering, mathematics or law. As shown in the introduction, for example, Sustainable Systems Design can be complementarily presented from the perspectives of systematic engineering design or mathematics. We regard the different forms of representation in ‘languages’ and types rather as a benefit than a disadvantage. This viewpoint is represented throughout this book.

3.2 Uncertainty Propagation Through Process Chains

Peter F. Pelz

We consider process chains, systems and structures in all product life phases in view, cf. Fig. 3.1. Therefore, it is obvious that we need an abstract representation in mind on how to treat those chains in a generic way. Figure 3.2 shows the model of a generic process chain being composed of single processes.

Representation of process chains

In Fig. 3.2 the physical product or system moves from left to right through the process chain, or the system or product is represented by the process chain itself. The evolution of the density function $p(\theta)$ of the product or system feature $\theta = \bar{\theta} + \delta\theta$ along the process chain is shown. In each process, energy, material and information are fed into or are withdrawn from the product. This is done by the work equipment, a generic term for man or machine. This incorporates tools such as sensors, actuators, controllers, and equally energy transformers and storage devices, i.e. capacities. Such a process may be active or semi-active, transient or quasi-stationary; it may

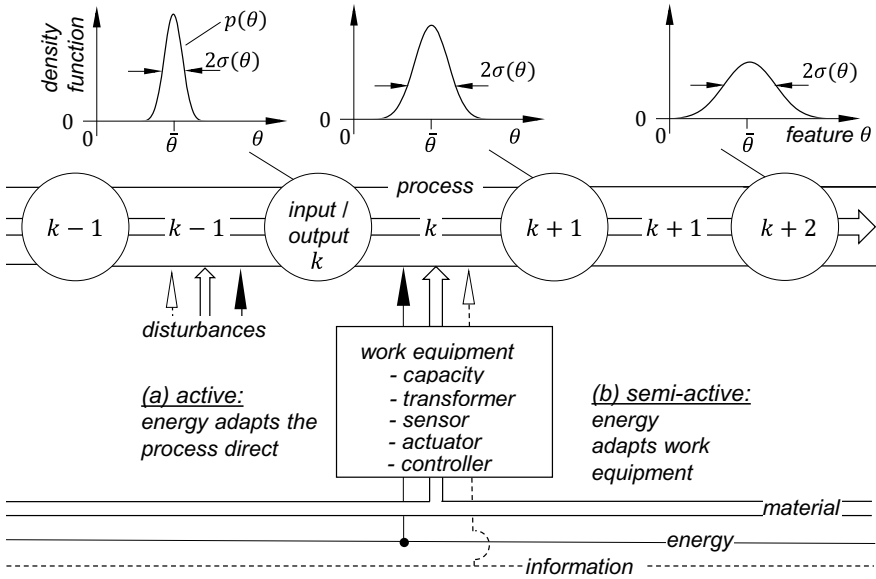


Fig. 3.2 Model of a generic process chain

be a static process, a relaxation process or a dynamic process. The different process characteristics will be discussed in Sect. 3.4.

In Chap. 1 we introduced engineering design thinking of structures built from components. This engineering design is represented by the generic process chain but also the branched structure or process chain depicted in Fig. 3.3. To model the general linkage of single processes, serial or parallel, the multi-pole presentation has spread [31]. The multi-pole representation is an object-oriented modelling approach. The representation is used in many domains, e.g. acoustics, electrical engineering, structural mechanics or fluid power. Other multi-domain model tools such as the bond graphs or Modelica are based on multi-pole representations.

Each component, i.e. each individual process, is easily modelled by a transfer matrix A_1, \dots, A_M with nonlinear algebraic entries. Using a multi-pole representation, the input data θ_k of the k th process are mapped to the output data θ_{k+1} : $\theta_k = A_k \theta_{k+1}$. Even though the reverse representation is often given, this representation is more concise when it comes to process chains. In a dynamical setting, the transfer matrices A_k may be given in the frequency domain. With a parallel arrangement of N components, the total transfer matrix is obtained by the sum of the individual matrices (i for input, o for output)

$$\theta_i = \theta_o \sum_{k=1}^N A_k.$$

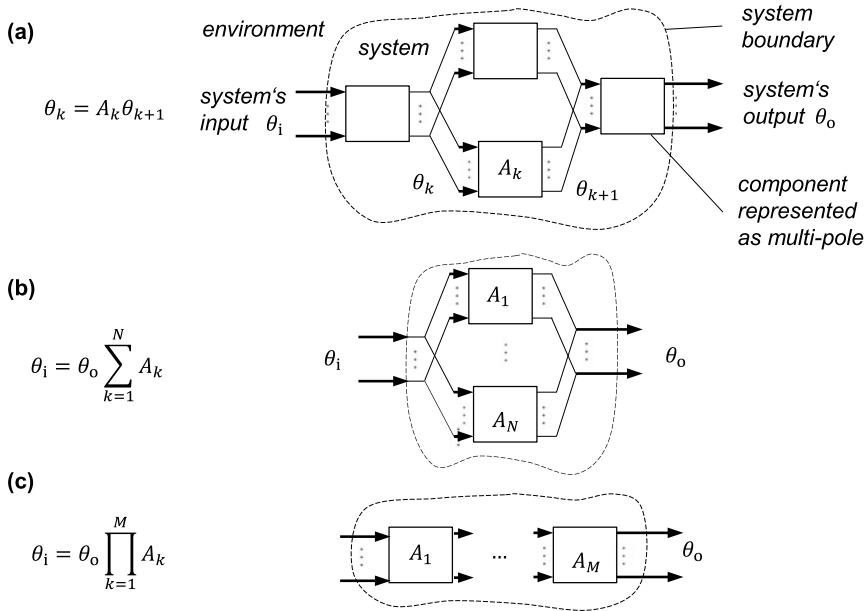


Fig. 3.3 a Multi-pole representation of a system or process chain; b parallel arrangement; c serial arrangement

With a serial arrangement of M components, the total transfer matrix is gained by the product of the individual matrices,

$$\theta_i = \theta_o \prod_{k=1}^M A_k.$$

Therefore, the model of the structure or the branched process chain can be easily assembled from the component models. Thus a model is available, which serves as a prerequisite for an uncertainty analysis or uncertainty propagation.

Uncertainty propagation

The process chains sketched in Figs. 3.2 and 3.3 help us to picture uncertainty propagation. The methods used in this book for uncertainty propagation are (i) moment methods and (ii) Monte Carlo simulations (MCS), cf. Sect. 2.1. The moment methods are the simplest and most commonly used methods. However, they are less powerful than the Monte Carlo methods. The latter are suitable in the preliminary design phase when compact, coarse granular models are used. In the later design stage, however, brute-force methods like MCS are often too expensive in terms of computing time.

In the following, we will shed some light on the Gaussian uncertainty propagation. This moment method is the most common one used by engineers, and it is worthwhile

to point out the conditions to be met when using this method. The experienced reader might skip the following section in small font. The case study presented in this section uses MCS in a preliminary design phase.

For the moment method, to be concise, we present the case of two processes 1 and 2 that are combined to an output $\theta_3 = \theta = h(\theta_1, \theta_2)$. The density function of $p(\theta)$ of the product or system feature θ is given by Sivia and Skilling [49]

$$p(\theta) = \int_0^\infty \int_0^\infty p(\theta_1, \theta_2, \theta) d\theta_1 d\theta_2 = \int_0^\infty \int_0^\infty p(\theta_1, \theta_2) \delta[\theta - h(\theta_1, \theta_2)] d\theta_1 d\theta_2.$$

This generalised convolution integral reads as follows: the density distribution function $p(\theta)$ of the process chain's output $\theta = h(\theta_1, \theta_2)$ is the integration of the joint density function $p(\theta_1, \theta_2)$ along the contour line $\theta = h(\theta_1, \theta_2)$. We consider the following important in the special case (i) and the general case (ii):

- (i) additive processes $\theta = \theta_1 + \theta_2$,
- (ii) Gaussian error propagation for $\theta = h(\theta_1, \theta_2)$.

Firstly (i), additive processes with $\theta = \theta_1 + \theta_2$ are widely used in mechanical engineering. One example of an additive process is the mounting of two bars each having a length θ_1 and θ_2 respectively. We may equally think of the mounting of two springs in parallel or two compliances in series. In the first case $\theta_{1,2}$ denotes the spring's stiffness. In the latter $\theta_{1,2}$ characterises the inverse stiffness, i.e. the compliance. We may even conceive a flow process in chemical engineering, mobility research or logistics. In this case $\theta_{1,2}$ denotes the residence time within each individual process and θ the resulting residence time.

Inserting $h(\theta_1, \theta_2) = \theta_1 + \theta_2$ into the generalised convolution integral using the product rule $p(\theta_1, \theta_2) = p_1(\theta_1)p_2(\theta_2)$ for statistically independent processes, yields the convolution integral for the additive process chain

$$p(\theta) = \int_0^\infty p_1(\theta_1)p_2(\theta - \theta_1) d\theta_1.$$

The density functions are normalised, i.e. the zero moment equals one. The first moment is the mean or expected value, the second central moment is the variance:

$$\bar{\theta} := \int_0^\infty \theta p(\theta) d\theta, \quad \sigma^2 := \int_0^\infty (\theta - \bar{\theta})^2 p(\theta) d\theta = \int_0^\infty (\delta\theta)^2 p(\theta) d\theta.$$

Using the convolution to express $p(\theta)$ in the two definitions for the mean and variance yields for additive processes the well-known result

$$\bar{\theta} = \bar{\theta}_1 + \bar{\theta}_2, \quad \sigma^2 = \sigma_1^2 + \sigma_2^2.$$

The derivation does not rely on Gaussian density functions, i.e. the result is exact for arbitrary density functions. The generalisation for more than two individual processes reads $\bar{\theta} = \sum \bar{\theta}_k$, $\sigma^2 = \sum \sigma_k^2$.

Recall the example of two assembled bars; it is trivial that the mean value of the total length is the sum of the mean values of the components. But also the variance of the length of the overall structure is the sum of the variances of the bars. For two flow processes in serial connection, the mean value of the total residence time is the sum of the mean residence times within the individual processes. Diffusive mixing can cause a sharp front to blur. Due to the linearity of the diffusive transport, the variance of the signal at the output is equal to the sum of the variances of the two individual processes.

The mentioned flow process is a nice visual model for the propagation of uncertainty in additive processes. This is due to the fact that the density distribution at the outlet $p(\theta)$ can be measured as a change in concentration over time in a single experiment: assume we would add N objects

(molecules, ...) in an arbitrarily short time at a time $t = 0$ into a constant flow Q . If the process has the volume V , then we measure the temporal change of the concentration given by $N/(V\bar{t}) p(t)$ at the process output. Here, \bar{t} is the mean residence time given by $\bar{t} = V/Q$. As always, things become clearer in a dimensionless representation: if we measure time in multiples of the mean residence time, i.e. $\theta := \frac{t}{\bar{t}}$, and concentration c in multiples of the typical concentration N/V , we measure $\delta(\theta) \rightarrow p(\theta)$, i.e. a Gaussian normal distribution for a purely diffusive process centred at $\theta = 1$. In a real experiment it is not possible to realise a Dirac delta function $\delta(\theta)$ but it is easy to realise a Heaviside function $\mathcal{H}(\theta)$. This is achieved by changing the concentration at the inlet at a time $\theta = 0$ suddenly from zero to one. In practice this means switching a valve. At the process outlet we then measure the cumulative density function $\mathcal{H}(\theta) \rightarrow P(\theta) = \int_0^\theta p(\tau) d\tau$.

Secondly (ii), in the general case, the dependent variable θ is a nonlinear function of one, two or more independent variables $\theta = h(\theta_1, \theta_2)$. A perturbation of h at the operation point $(\theta_{10}, \theta_{20})$ marked by the index 0 is given by

$$h = h_0 + \left. \frac{\partial h}{\partial \theta_1} \right|_0 \delta\theta_1 + \left. \frac{\partial h}{\partial \theta_2} \right|_0 \delta\theta_2 + \mathcal{O}(\delta\theta^2).$$

Inserting this perturbation into the general convolution integral and neglecting terms of the order of magnitude $\mathcal{O}(\delta\theta^2)$ yields the Gaussian uncertainty propagation [29]

$$\sigma_g^2 \approx \left[\left. \frac{\partial h}{\partial \theta_1} \right|_0 \right]^2 \sigma_1^2 + \left[\left. \frac{\partial h}{\partial \theta_1} \frac{\partial h}{\partial \theta_2} \right|_0 \right] \sigma_{12}^2 + \left[\left. \frac{\partial h}{\partial \theta_2} \right|_0 \right]^2 \sigma_2^2.$$

The index g indicates the Gaussian approximation of the variance. Here,

$$\sigma_{12}^2 := \int_0^\infty \int_0^\infty (\theta_1 - \bar{\theta}_1)(\theta_2 - \bar{\theta}_2) h(\theta_1, \theta_2) d\theta_1 d\theta_2 = \overline{\theta_1 \theta_2} - \bar{\theta}_1 \bar{\theta}_2$$

denotes the co-variance of the two independent variables.

The Gaussian uncertainty propagation is the simplest and most commonly used method of uncertainty propagation, but it is rarely critically questioned. From the derivation it is clear that the trust into the method should decrease with increasing nonlinearity in $h(\theta_1, \theta_2)$. This becomes clear for the simpler case $\theta = h(\theta_1)$. This coordinate transformation shown in Fig. 3.4b leads to a distortion of a symmetric input density function $p_1(\theta_1) \rightarrow p(\theta)$. As a result, the mean $\bar{\theta} = \int_0^\infty h(\theta_1) p_1(\theta_1) d\theta_1$ differs from the approximation $\bar{\theta}_g = h(\theta_{10})$ and further σ_g^2 is only an approximation of σ^2 , Fig. 3.4b. In summary we recognise that the usage of the Gaussian approximation $(\delta\theta)^2 \approx (\delta\theta)_g^2 = \sum (\partial h / \partial \theta_k)^2 (\delta\theta_k)^2$ for uncorrelated input data θ_k needs some justification.

For the strongly nonlinear process chain being treated in the following Sect. 3.3, the moment method, i.e. the Gaussian uncertainty propagation is not adequate. In Sect. 3.3 we use the bootstrapping method [10]. Other methods and examples for uncertainty propagation are found in Sect. 6.1.6.

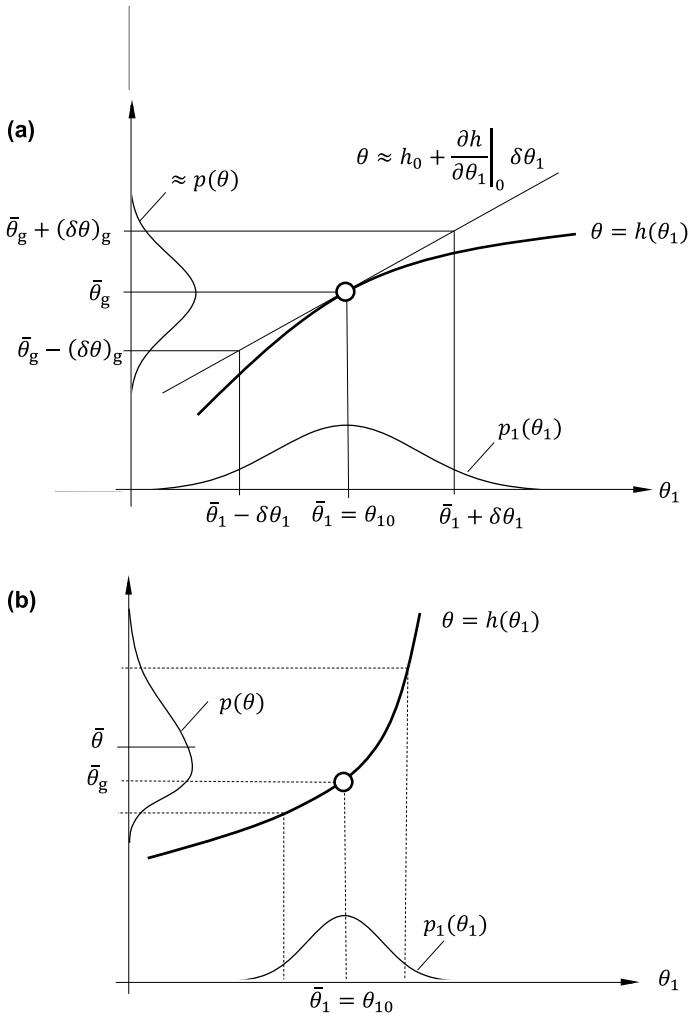


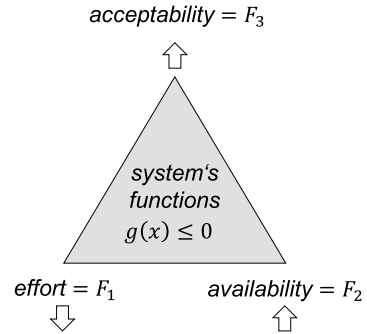
Fig. 3.4 **a** Gaussian uncertainty propagation shown to be associated to a coordinate transformation or mapping $\theta_1 \rightarrow \theta = h(\theta_1)$; **b** approximation $\bar{\theta}_g \approx \bar{\theta}$ due to the nonlinearity in $h(\theta_1)$

3.3 Five Complementary Methods for Mastering Uncertainty in Process Chains

Peter F. Pelz

In Sect. 3.2 we applied moment methods of uncertainty propagation to process chains. In several parts of the book we resume these and use more in-depth methods of uncertainty propagation, cf. Sect. 6.1.6. Uncertainty quantification includes uncertainty propagation [30]. Uncertainty propagation is often recognised as uncertainty

Fig. 3.5 Sustainable Systems Design: Optimal quality subject to functionality



analysis. However, it is only one of at least five methods that form a complementary toolbox for the mastering of uncertainty in mechanical engineering.

We will discuss the five methods of the toolbox from the perspective of Sustainable Systems Design, which was presented in Chap. 1: the shift to the new paradigm ‘optimal quality subject to functionality’ combines the two incomplete paradigms ‘form follows function’ (Sullivan) [51] and ‘less but better’ (Rams) [42]. The quality $F = \{F_1, F_2, F_3\}$ is represented by (i) the effort F_1 to manufacture and operate the system, (ii) the availability F_2 of the system, as well as (iii) the acceptability F_3 of the system by the user and the society. The functionality is represented by the system’s functions, which are represented by a subset of the constraints, $g(x) \leq 0$. Here, functional quality is one aspect of acceptability and is measured by the deviation δg from the customer expectation given by the specifications g_s . The lower the loss in functional quality δg , the higher the acceptability.

In summary, this understanding of the interplay between function and quality as described here as well as in Sect. 1.7 and Fig. 3.5 is more general than the understanding in the concept of Robust Design [41]. With Robust Design, firstly, effort is measured in costs, secondly, availability is understood as part of the function, and thirdly, the only measure for acceptability is functional quality. As pointed out in Sect. 1.7, the more general understanding of function and quality allows the mapping of Sustainable Systems Design to a constrained optimisation problem, provided the metrics for effort F_1 , availability F_2 and acceptability F_3 are well defined, cf. Fig. 1.10.

In this section, we will introduce the methodological building blocks and show the interaction of the methods by means of a use case. The five methods addressed here are

- (i) uncertainty identification,
- (ii) uncertainty propagation,
- (iii) robust optimisation under uncertainty and/or multi-objective optimisation,
- (iv) sensitivity analysis,
- (v) model adaptation.

Firstly (i), for uncertainty identification we developed the concept of “data-induced conflicts”. Here, the trust in underlying data sources is identified, cf. Sect. 4.2. Sec-

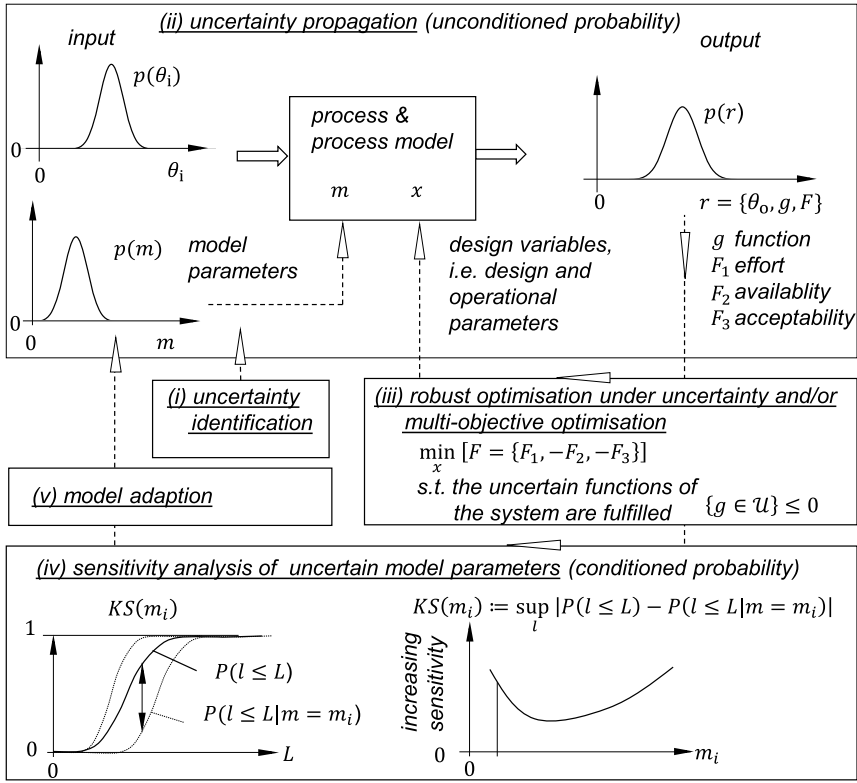


Fig. 3.6 Relation between (i) uncertainty identification, (ii) uncertainty propagation, (iii) optimisation under uncertainty, (iv) sensitivity analysis of uncertain model parameters, (v) model adaption. Inputs are θ_i and design variables x , outputs are θ_o , which depend on the constraint functions g and objectives F_1, F_2, F_3

only (ii), uncertainty propagation helps us to quantify unconditioned uncertainty of system output data z , system functions g and system objectives F , cf. Sect. 4.3. Thirdly (iii), robust optimisation under uncertainty combined with multi-objective optimisation helps us to select the design variants x robustly, i.e. in such a way that the system quality F is optimised subject to the uncertain system function g . In other words, the system has an optimal quality F subject to the condition that it has to satisfy not only one, but a set of constraints in the uncertainty set $g \in \mathcal{U}$. Fourthly (iv), sensitivity analysis allows us to indicate those model parameters m which are most relevant for model adaption (v).

The interaction of the five methods with a generic process and the data types introduced in Chap. 2 is best understood in a control loop or flow diagram with the generic input-output-process model, shown in Fig. 3.6 in its centre. At the boundary of the generic process, we recognise the data introduced in Chap. 2, i.e.

- (i) functional requirements g and quality data $F = \{F_1, F_2, F_3\}$,

- (ii) input u split into design variables x and other input variables,
- (iii) output z ,
- (iv) model parameters m , e.g. mass flux, energy flux, information,
- (v) internal variables y .

The overall data of the generic process subject to uncertainty are collected into $r = \{\theta, g, F\}$. The model parameters m and design variants x are inherent to the model as depicted in the flow diagram. Perturbations to the model are not shown but recognised as present. Note that in terms of uncertainty propagation in process chains of Sect. 3.2 the input would be $\theta_i = \theta$ (minus the design variables) and the output $\theta_o = z$. Moreover, the assignment of data to a certain type is not strict. For example, it is common in optimisation to formulate a sub-objective as a constraint, cf. Sect. 5.1.1. Depending on the context and the method, a model parameter can also become an output datum, for example.

To show the relations between the five methods (i)–(v) we will go clockwise through the control loop starting from the generic process. The reliability of underlying data sources (i) is evaluated by means of the concept of data-induced conflicts presented in Sect. 4.2.

The result of the generic process comprises the data z , given by the flows of mass, energy and information, further the system functions g and the system quality F . The density function $p(r)$ is the result of uncertainty propagation (ii) through the process. The propagation of incertitude or stochastic uncertainty through general nonlinear processes are easily applicable for engineers and Monte Carlo experiments form a practicable method. The interpretation of the Monte Carlo experiments is based on the “law of large numbers”, see Chap. 2. That is, the signal to noise ratio of the experiment outcomes increases with the number of calculated samples. Hence, the downside of the brute force method in contrast to the moment methods can be the high computational effort.

The result of the robust optimisation under uncertainty combined with multi-objective optimisation (iii) are suggestions how to select the design variants x . Depending on the subjectively selected weights between the three different sub-objectives, the components are selected, their size is determined provided the selection is done from a size-ranked modular design and finally the operational conditions are determined. Hence, the optimal design and operational variants are determined by the robust optimisation under uncertainty and/or multi-objective optimisation.

Sensitivity analysis (iv) is the fourth method in the control loop, cf. Fig. 3.6. In addition to the input data w and the design variants x , there are predefined model parameters m . In axiomatic models, these are often material parameters of constitutive equations, such as the modulus of elasticity in Hooke’s law. However, model parameters can also be dimensionless parameters as mentioned above.

Iooss et al. [30] point out that “sensitivity analysis provides users of mathematical and simulation models with tools to appreciate the dependency of the model output from the model input and to investigate how important each model input is in determining its output”. There are many methods for sensitivity analysis. Saltelli et al. [44] provide a concise review paper on this topic. Also Ghanem et al. [15] give

an overview on sensitivity analysis in Part IV of their “Handbook of Uncertainty Quantification”: there are local and global sensitivity methods. The local methods are gradient based, whereas most of the global sensitivity methods are probabilistic. For the method considered here, in contrast to uncertainty propagation, probabilistic sensitivity analysis is based on Bayes theorem, cf. Sect. 4.1.2, i.e. on the concept of conditional probability. From the large number of local and global methods for sensitivity analysis we present only one. A very general, efficient and hence for engineers easily applicable global sensitivity method is the PAWN method used by Holl et al. [27].

From all model parameters m one is selected, i.e. m_i . The sensitivity with respect to the selected parameter is given by Kolmogorov–Smirnov metric as originally proposed by Baucells and Boronovo [3]. This is the largest distance

$$KS(m_i) := \sup_l |P(l \leq L) - P(l \leq L | m = m_i)|$$

between the unconditional cumulative probability function $P(l \leq L)$ and the large number of conditional cumulative probability functions $P(l \leq L | m = m_i)$. Here the model parameter $m_i \in [\min, \max]$ is selected arbitrarily from its interval. We have L numerical experiments. The l th experiment is identified by the numerator l .

The larger the sensitivity of m_i the more attention is to be paid to this parameter. The model adaption (v) is gained manually or again by means of optimisation methods. The model adjustment can have several stages: Firstly, model parameters can be adapted. Secondly, the model granularity can be adapted and thirdly, the model type or structure can be customised. The adaptation of model parameters mentioned here is mostly understood as calibration. The model granularity concerns the time and space resolution. The most profound change is the adaptation of the model structure. Here, model structure should not be mismatched with structural uncertainty introduced in Sect. 2.3.

In a data-driven model, the structure is given e.g. by the number of layers in an artificial neural network and the calibration determines the weights of each neuron’s input. Obviously the change of structure is more profound than the calibration of the weights, and obviously the calibration is based on optimisation. The counterpart of the data-driven model is the axiomatic model. The structure of such a model must be adjusted if the model is to describe wave propagation, however, the model is represented by an elliptic differential equation. The calibration of data-driven models is widely used today. More interesting is the structural adaptation of axiomatic or grey-box models.

Four case studies are included in the chapter to illustrate the content. Depending on the reader’s interests, the case studies may be read or skipped.

The first case study presented applies the concept of a multi-pole process chain, uncertainty propagation by MCS and sensitivity analysis. Furthermore, firstly the concept of parameter optimisation based on scalable modules is presented, and secondly the transformation of a transient investment into a time-invariant cost flow is shown.

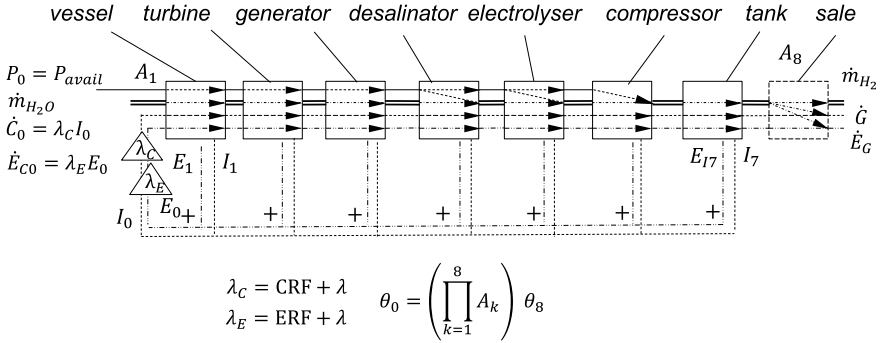


Fig. 3.7 Multi-pole representation of the energy ship; the structure is predefined; the size of the components, i.e. wind powered vessel, turbine, generator, desalinator, electrolyser, compressor and tank are design variants; the same is applicable to the operating parameters, such as the rotational speed of the turbine; to evaluate even ecological gains one information stream may equally be a flux of ecological costs \dot{E}_C

1st case study—Mastering uncertainty in the earliest design phase

By way of illustration, we show how the named methods can be applied to control uncertainties in the earliest design phase of a regenerative system. The concept of an ‘energy ship’ is to be designed and evaluated under uncertainty. This ship is supposed to convert wind energy into electric energy using a wind-powered vessel. A hydrokinetic turbine integrated into the vessel is used to convert kinetic energy of the water into electricity. In a further step, the electricity is used to produce hydrogen using sea water.

The obvious advantages of the ‘energy ship’ compared to far off-shore wind power are threefold: firstly, the energy supply system solves the transport problem of off-shore electricity production inherent in the system; secondly, availability is increased compared to existing solutions by allowing the ship to avoid a storm; thirdly, known modules, each with known uncertainty, are combined to form a new system. We will return to this point in Sect. 3.5 when speaking about flexibility enabled by combining modules.

The process chain needed as a basis for a decision for or against a technology made under uncertainty is a techno-economic model of the ‘energy ship’ using the multi-pole representation introduced in Fig. 3.7. The fluxes through the chain are the fluxes of energy, mass and money. The system model is formed by the process chain, the model uncertainty is given by the uncertain physical and economic parameters.

To shed some light on model uncertainty, here, we will only focus on the economic model. The physical model parts may be found in [40]. The simple balance reads profit=revenue-costs: $\dot{G} = \dot{R} - \dot{C}$. There was no periodic profit \dot{G} , if the mass specific production cost of hydrogen would equal the market price f_{H_2} . Hence, the system design should be oriented at the threshold $\dot{G} = 0$. If the production costs equal the market price, the production costs are named levelised cost of hydrogen (LCOH), i.e. $\dot{G} = 0 : LCOH = f_{H_2}$.

The main and first part of the costs are due to the investment costs I_0 . The smaller second part of the costs are due to the maintenance; usually they are assumed to be proportional I_0 . The factor λ is a rate, i.e. the dimension is an inverse time. Hence, the maintenance costs rate are λI_0 .

Since I_0 usually has to be spent immediately at the beginning of the project, investment is inherently a time variant process given by Dirac’s delta function, $I_0 \delta(t)$. In order to make decisions under uncertainty, time-variant problems are much more difficult to treat than time-invariant problems, cf. Fig. 3.11. Fortunately, we can transform the transient investment $I_0 \delta(t)$ into an equivalent

time-invariant cost-flux $\dot{C}_0 = \text{const}$ using a method developed by Simon Stevin as early as 1582 [50]: the time-variant process of increasing money volume for a fixed value or changing value of a fixed amount of money is captured by the interest rate minus the inflation rate, i.e. the effective interest rate $z(t)$. The time-variant process of the information data “money” is given by the evolution equation

$$\frac{dZ}{dt} = z(t)Z, \text{ with } Z(0) = I_0.$$

Here, we use the shorthand and much clearer notation of calculus instead of the usual summation notation. The evolution equation represents the development of money volume on the money market due to compound interest. Solving this simple initial value problem using the transformation to a dimensionless time $\tau(t) = \int_0^t z(t') dt'$ leads to $Z(t) = I_0 e^{\tau(t)}$. Now, I_0 is not put on the money market at time $t = 0$ but invested causing a cost flux \dot{C} at time $t > 0$. The compound interest of the time span $[0, t]$ is covered by the factor e^τ . Hence, in determining the equivalent cost flux \dot{C} we demand the equivalent $dI_0 e^{\tau(t)} = \dot{C} dt$. Integration over the time period $[0, T]$ leads to $I_0 = \int_0^T \dot{C} e^{-\tau(t)} dt$. For the special case of constant cost flux $\dot{C} = \dot{C}_0 = \text{const}$ we can define the ratio of cost flux C_0 and invest I_0 as the capital recovery factor.

$$\text{CRF} := \frac{\dot{C}_0}{I_0} = \frac{1}{\int_0^T e^{-\tau(t)} dt} = \frac{1}{T} \frac{1}{1 - e^{-zT}}.$$

In analogy to λ , the dimension of CRF is a rate, i.e. $1/s$. The right side of the above equation is gained for constant effective interest rate $z = \text{const}$. As a result, the cost flux is given by the investment costs I_0 multiplied by the sum of the capital recovery factor CRF and the maintenance rate λ : $\dot{C} = I_0(\text{CRF} + \lambda)$. Hence, by Stevin’s “trick” we transformed a time-variant process into a time-invariant process.

On the other side, the revenue flux $\dot{R} = \delta T \dot{m}_{\text{H}_2} f_{\text{H}_2}$, gained by the sale of hydrogen, embodies the counterpart of the cost flux. Here, the capacity factor is denoted as δ . Hence, the dimensionless product δT measures the availability of the system introduced in Chap. 1.

From the balance $\dot{G} = \dot{R} - \dot{C} = \delta T \dot{m}_{\text{H}_2} f_{\text{H}_2} - I_0 (\text{CRF} + \lambda)$ we derive the levelised cost of hydrogen (LCOH) to be

$$\dot{G} = 0: \text{LCOH} = \frac{I_0(\text{CRF} + \lambda)}{\delta T \dot{m}_{\text{H}_2}}.$$

Here, \dot{m}_{H_2} is the hydrogen mass flow. Speaking again in system’s function and quality: the function of the system being here in focus is to convert energy. The quality of the system is given by LCOH. This quality is increased, if the effort measured in investment cost $I_0 = \sum_{k=1}^8 I_k$ is decreased, if the availability measured in δT is increased or if the maintenance cost rate λI_0 is reduced.

Hence, the design task is described by

$$\min_x \text{LCOH s.t. } g(x) \leq 0.$$

The constrains $g(x)$ are given by the mult-pole model of the system. There are binary and real decision variants. A binary decision is needed in deciding whether the hydrogen production on the ship uses salty seawater or stored fresh water. In the first case a desalinator is required as an additional module, in the latter case not; but using fresh water would require an additional tank with a limited capacity. The real design variants determine on the one hand the size of the modules. These are usually obtained as size ranged modules. On the other hand they determine the operational variant such as rotational speed of the turbine.

Here, as opposed to the example in Sect. 1.5 as well as further examples within this book, we ignore the binary decisions. Thus, in our example the structure of the system is predefined and sketched as the process chain in Fig. 3.7. The design variants are provided by the questions “what should be the length $x_1 = L$ of the vessel?”; “what should be the size of the turbine measured in a cross-sectional area $x_2 = A$?” etc.

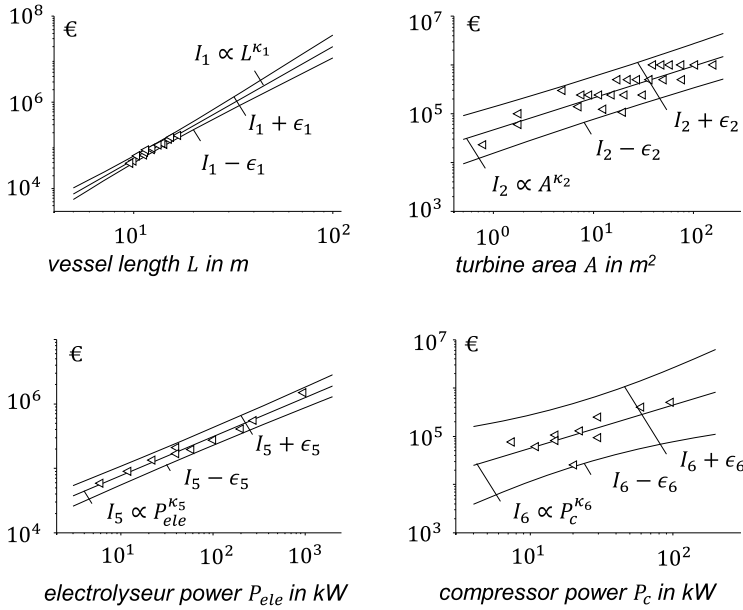


Fig. 3.8 Empirically derived investment cost models $I_k \propto x_k^{\kappa_k}$ for some of the modules needed to compose the energy ship; the 95% prediction interval $I_k \pm \epsilon_k$ is given [27]

The individual investment costs I_k scale with the design variants by means of power laws $I_k \propto x_k^{\kappa_k}$ as a market survey for size ranged components indeed shows, cf. Fig. 3.8.

Together with the physical model based on the energy, continuity and momentum equation, the techno-economical model is defined. The techno-economic-ecologic system model is a multi-pole model as described in the previous section. The model maps the input data $\theta_i = \theta_0$ to the output data $\theta_o = \theta_8$ by the matrix multiplication $\theta_i = A\theta_o$: $A = \prod_{k=1}^8 A_k$. The entries of the 4×4 system matrix are the technical, economic and ecologic models being described by Holl et al. [27].

Figure 3.9 shows that the techno-economically optimal system analysis is defined by a vessel length of 37 m with the help of stochastic optimisation. At the optimum, the turbine area is 0.9% of the sail area. For this and all the other design variants not shown here, the levelised costs of hydrogen are 13.9 €/kg. Comparing this quality datum with the 2–3 €/kg for methane reforming shows that the production costs, even for the optimal system, are still high.

So far we have only addressed the prediction uncertainty of the component models, cf. Fig. 3.7. A global sensitivity analysis allows us to identify those model parameters $m_i \in m$ on which the optimisation result shows the largest sensitivity. In return, when adapting the model one should concentrate on mastering the uncertainty in the most sensitive model parameter.

Figure 3.10 shows the mean of $KS(m_i)$ [3] for some model parameter. The result is most sensitive with respect to uncertainty in the sail’s lift parameter c_l . The system shows only a small sensitivity to the vessel displacement parameter relating the displacement volume to the cube of the ship’s length, $\epsilon_D := \frac{V}{L^3}$. Interestingly, the system shows only an intermediate sensitivity to the availability δT .

In summary, this section reveals that the process chain model is at the heart of methods for mastering uncertainty. The methods discussed are complimentary and differ, for example, in terms of conditional and unconditional probability. Finally, we have used an example to show how a system can be evaluated in an early design phase. Here,

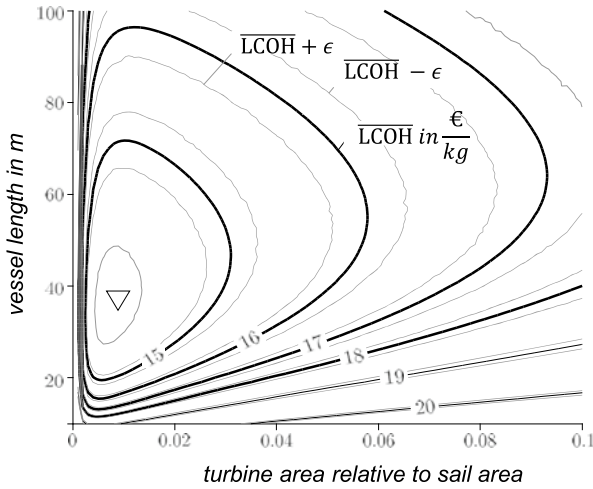


Fig. 3.9 Arithmetic average $\overline{\text{LCOH}} \pm \epsilon$ (ϵ represents 95%-confidence interval) for all combinations of dimensionless turbine area and vessel length; the sweet point is shown to be a triangle [27]

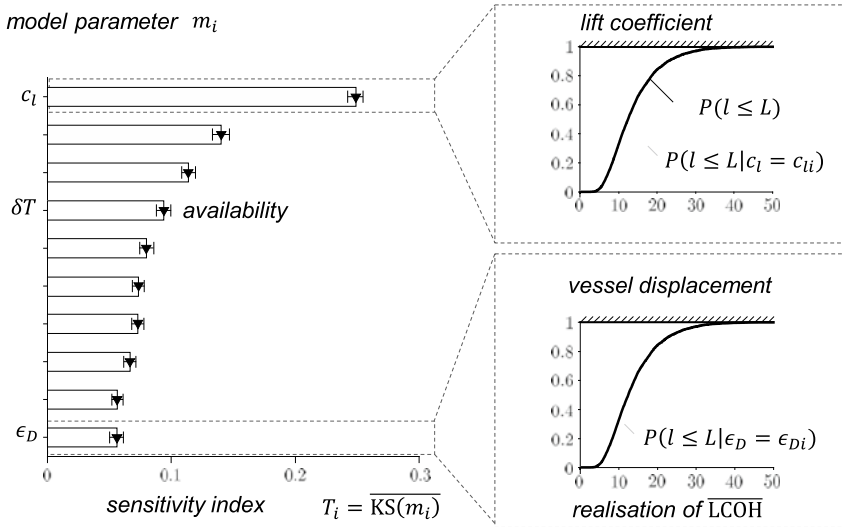


Fig. 3.10 Sensitivity analysis of the techno-economic optimal system (left) and graphical clarification on the example of the most sensitive and the less sensitive parameter [27]

it is necessary to map a time-variant process into a time-invariant process. Whether and how such a transformation succeeds is always problem- and model-specific. In the following section we will see how degradation can be treated in time-invariant terms, section by section. This leads us to the next section where we classify the time characteristics of process chains.

3.4 Time-Variant, Dynamic and Active Processes

Peter F. Pelz

Mechanical engineers are trained to solve problems in the field of dynamics, but they are much less familiar with the task of creating an optimal design of transient processes under uncertainty. Experience has shown that making decisions under uncertainty is much more difficult for time-variant processes than for time-invariant processes. Therefore, in the following criteria are to be compiled to determine when an actually time-variant process can be treated as a time-invariant process. Before doing this, it is worthwhile to classify the time characteristics of the processes.

Figure 3.11 shows the Euler diagram that classifies the time characteristics of processes. Synonymous for time-invariant is stationary, synonymous for time-variant is transient.

A process is quasi-stationary if it reacts immediately to a temporal change in the boundary condition. In dynamics, a process is quasi-static and at the same time quasi-stationary when the system is excited by a frequency f far below the system's lowest natural frequency: $f \ll 1/\sqrt{LC}$. Here L is the inductance representing a solid or liquid body or an electric coil and C is the capacity or compliance associated with the storage of energy in a spring, in a gas volume or within an electric accumulator.

There is a slightly different case when the external load is in balance with frictional, i.e. dissipative $\propto R$, and elastic, i.e. conservative $\propto C^{-1}$, forces. The system then shows a relaxation time which is given by the product RC . The relaxation time is the typical time a system needs for relaxing to an equilibrium when disturbed. For $f \ll (RC)^{-1}$ the system behaves quasi-stationary. The model granularity with respect to temporal resolution is determined by the functional quality of the process

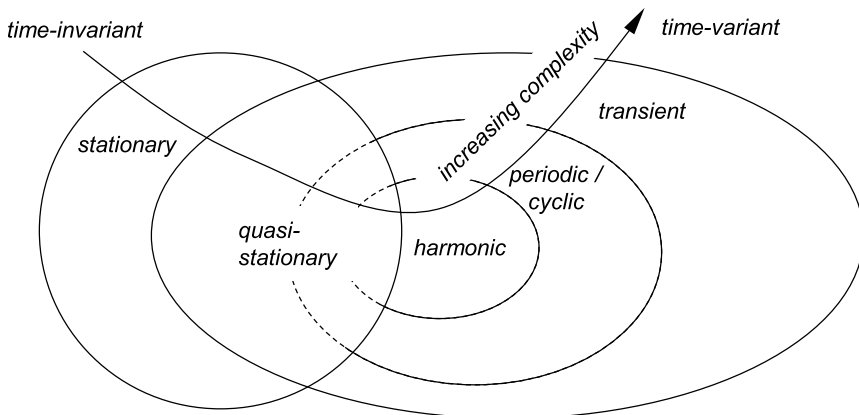


Fig. 3.11 Euler diagram for classification of time-variant and time-invariant processes

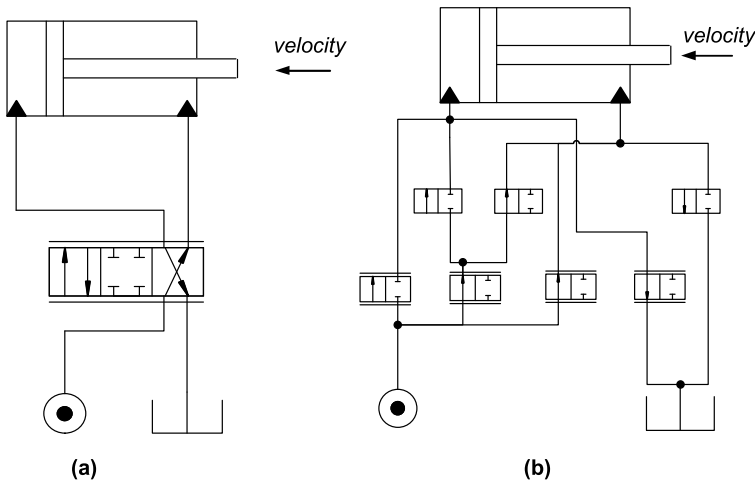


Fig. 3.12 Design of a hydrostatic transmission aiming for (a) the minimal number of valves; (b) minimal particle erosion over the life time of the hydrostatic transmission [53]

which shall be resolved in time. Thus the highest natural frequency being resolved is determined by the quality promise given explicitly or implicitly to the customer.

Special types of time-variant processes, i.e. transient processes, are periodic, i.e. cyclic processes. Again, a special type of cyclic processes are harmonic processes. Figure 3.11 indicates that the process complexity in terms of time behaviour is reduced if a time-variant process can be mapped to a time-invariant process or if a transient process can be mapped to a harmonic process. An example of the former was given in the first case study, in which the time-variant investment was mapped to the equivalent time-invariant cost flux, for the latter transforming the linearity of the process is usually required as a prerequisite.

The second case study presented now is a degradation process, i.e. a time-variant relaxation process. The lesson learned from this will be how degradation can be treated for system design under uncertainty.

2nd case study—Modelling of degradation

Figure 3.12 shows two different optimal hydraulic structures (a) and (b) a hydrostatic transmission that performs the very same function. The function is given by a load profile. This is cyclic with phases of constant and possible uncertain load and speed. For an assumed ideal rigid system, i.e. vanishing compliance $C \rightarrow 0$, the system behaves quasi-stationary with respect to power transmission. For the algorithmic system design it is allowed to select 2/2-way and 4/3-way proportional and non-proportional valves.

For structure (a), the design objective has been to minimise the number of valves needed to establish the hydrostatic transmission. For structure (b), the design objective has been to minimise the particle wear of the valves in each load cycle over the life time of the hydrostatic transmission. Mastering the structural uncertainty, cf. Sect. 2.3 in case (a) is simple: the design solution with two proportional and two non-proportional valves is obvious. The control of the structural uncertainty

in case (b), on the other hand, is far from obvious. Although the system is dynamically time-invariant, due to the degradation process the system is time-variant in a long time frame. To master the structural uncertainty in the design process a degradation model has to be envisaged in the constrained optimisation.

In this case study, the degradation is depicted very generally as a relaxation process. This approach can be applied to many other ageing and wearing processes: particle erosion is one possible reason for degradation. Others are chemical degradation or hardening of polymer material due to temperature activated diffusion and/or chemical reaction. Also damage accumulation in cyclic loading of components may be treated as a degradation process.

For each of the mentioned processes, the degradation of a component property F may be described by an evolution equation $\frac{dF}{dt} = f(t, L(t), F)$. Here, t is the time and $L(t)$ the load history of the component. The initial condition is $F(t = 0) = 1$ for the degradation process and $F(0) = 0$ for the complementary accumulation process. A separation approach of the right hand side is a good model of most degeneration or accumulation processes seen in nature and technology.

$$\frac{dF}{dt} = K(L(t))f(F). \quad (3.1)$$

Here, the load history $L(t)$ and the time t enter the equation only implicitly by means of the degradation rate $K = K(L)$. The dependency is implicit, since only the rate depends on the load history. In fact we came across the evolution equation in the previous Sect. 3.3 where the rate is represented by the interest rate $z(t)$. The interest rate is determined by the global market and the central bank. The degradation rate is determined by the process intensity, i.e. the load. If we know $K(L)$ and the load history $L(t)$ we know the erosion process. The load history $L(t)$ is either determined by the measurement in the usage phase or by providing a model for the expected load history $L(t)$ based on our past experiences gained in the same or a similar process.

Transformation to a dimensionless time $d\theta := K dt$, $\theta(0) = 0$, yields $\frac{dF}{d\theta} = f(F)$. Thus, with an increasing load the rate K increases and the physical time t moves faster. Here, $F := M/\widehat{M}$ is a product property M in relation to its initial value \widehat{M} in the case of a degeneration or to the saturation value \widehat{M} in the case of an accumulation influencing the product's function.

In technical systems often the initial degradation or accumulation phase is more relevant than the later phase. With a view to the asymptotic limit in the initial phase, the initial phase on the right-hand side of the evolution equation is independent of F . Hence, an approximation for the initial degradation or initial accumulation process is given by $\frac{dF}{d\theta} \approx \mp 1$. This leads to the damage accumulation.

$$\frac{M(t)}{\widehat{M}} = \theta(t) = \int_0^t K[L(t')] dt', \quad (3.2)$$

known as Palmgren–Miner's rule [33, 37].

Figure 3.12b shows the optimal hydraulic structure with regard to the minimal particle erosion over the lifetime of the hydrostatic transmission. For particle erosion, the erosion rate is a function of the volume fraction c_v of solid particles within the fluid, the pressure drop history $\Delta p(t)$ across the valve, the history of the relative valve opening $0 \leq z_+(t) := \frac{z(t)}{z_{\max}} \leq 1$, the maximal cross sectional area $\pi D z_{\max}$ being proportional to the diameter D of the spool valve and the oil density ϱ , cf. Fig. 3.13. For dimension reasons we have the equivalent representation of the degradation rate $K(c_v, z_+, \Delta p, \varrho, \pi D z_{\max}) = K_+(c_v, z_+(t))\sqrt{\Delta p(t)}/(\pi D z_{\max} \varrho)$. For a load profile such as the one sketched in Fig. 1.7b the degradation process may be treated as stationary in each load phase of the load cycle.

The valve's function is given by the pressure amplification factor $V := \frac{dp}{dz}$. Hence, the functional degradation Δz in each stationary phase of duration Δt is given by

$$\frac{\Delta z}{z} = \frac{\Delta z_+}{z_+} = \exp\left(-\frac{K_+}{3} \sqrt{\frac{\Delta p \Delta t^2}{\pi D z_{\max} \varrho}}\right) - 1. \quad (3.3)$$

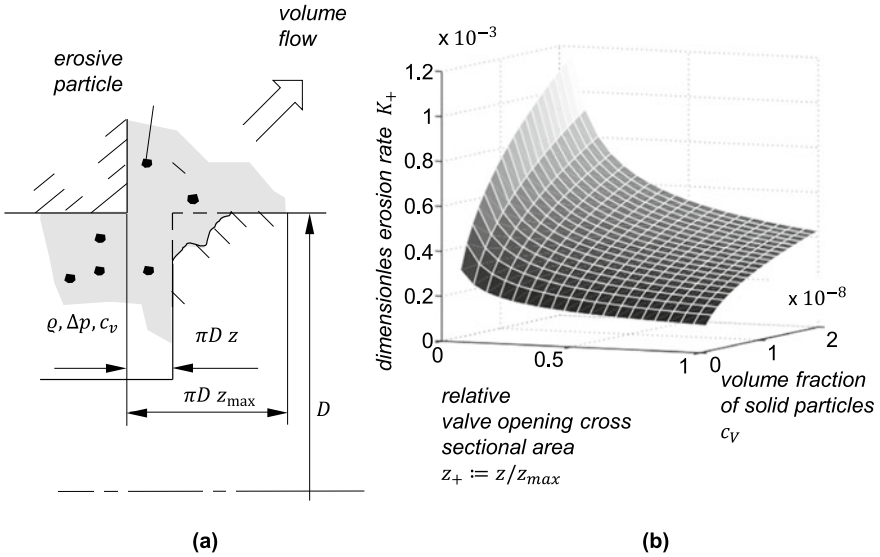


Fig. 3.13 Erosion model of the particle wear of a spool valve **a**. The experimental data **b** are generalised by means of a dimensional analysis [53]

With a view to finding an optimal structure in terms of minimal wear depicted in Fig. 3.12b a mixed integer nonlinear program (MINLP) is to be solved by

$$\min_x \max_{\text{valves}} \sum_{\text{load phases}} \Delta z \text{ s.t. } g(x) \leq 0. \tag{3.4}$$

In the MINLP the hydrostatic transmission is represented by a complete graph. Each possible valve, Fig. 1.7a, is mapped to an edge of the graph. The nodes are the hydraulic connections. The constraints $g(x)$ are given on the one side by the continuity and energy equations for each edge or node. On the other side they are given by the load profile sketched schematically in Fig. 1.7b.

As lessons learned, firstly, mastering structural uncertainty requires here the solution of a MINLP. Secondly, model uncertainty is relevant for the degradation model depicted in this case study. Thirdly and finally, the load profile sketched in Fig. 1.7b may be uncertain itself. Taking into account this uncertainty requires robust optimisation, cf. Chap. 6 techniques as presented in this book.

Costs and gains of active components

If one thinks of technologies to master uncertainty, one can think of active or “smart” technologies. Active technologies promise new degrees of freedom and the ability to adapt to unforeseen circumstances. Despite this promise, it is surprising to find that in practical applications far fewer active solutions are implemented than engineering research would suggest.

On the one hand, a reason for this is that the total costs for integrating active components can be considerable. On the other hand,—contrary to the actual intention—active components can also make a system vulnerable, because active components do not only include an actuator but also sensors, controllers as embedded software modules, software interfaces, cables, and connectors, all of which may fail. As seen in Chap. 1, training, legal conformity certification, commissioning, release, versioning, maintenance, procurement and much more increase the number of topics that influence the uncertainty of the system from a holistic point of view. If these are not mastered, the system may possibly be damaged.

It is therefore a conflict of objectives that has to be resolved. What are the costs and gains of the active components? Figure 1.11 helps with orientation: it is a matter of clearly formulating and weighing the function and the objectives and then deciding whether to use active components or not. What is also helpful is to answer the question: “What does an active component promise in an ideal case?”. The answer to this question is an answer such as “as good as it gets” with respect to effort, availability, acceptability. The answer to the question “What does an active component promise in an ideal case?” usually leads to a Pareto boundary, Sect. 1.6.

Also here, mastering uncertainty is made easier if individual processes are defined. Hence, it is worthwhile to classify types of active processes. In the generic process chain shown in Fig. 3.1 we distinguish between active and semi-active processes. The classification is as follows [6]:

- (i) for an active process, energy is transferred to the process and finally to the product or system directly;
- (ii) by a semi-active process only the work equipment is adapted by the input energy;
- (iii) a passive component or process is neither active nor semi-active.

A passive process may be a storage or a transport process. A passive component is e.g. a coil spring. An illustrating example of a semi-active process is the semi-active damper shown in Fig. 1.2. Manually, the damper setting may be changed from “hard” to “soft” by rotating the actuating lever. In the usage phase, no external energy is supplied for controlling the two connected sub-systems. This is different for the Active Air Spring being presented in Sect. 3.6.2 where the external energy is used to isolate two dynamic systems from each other. Thus, even the system’s resonance when excited at the natural frequency may be suppressed by the active system.

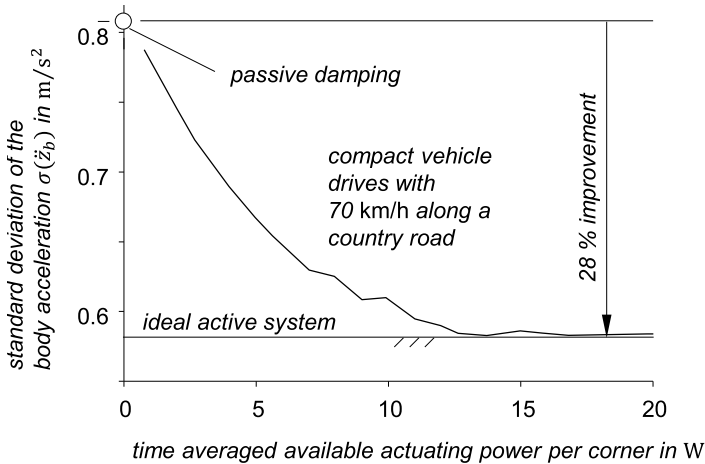


Fig. 3.14 Improvement of functional quality through an active component, here the Active Air Spring: A compact vehicle drives over a country road at a constant speed of 70 km/h. The standard deviation of the body acceleration is plotted against the time-averaged actuating power. The horizontal asymptote is determined analytically under the assumption of unlimited actuating power and actuating speed [43]

Example—Suspension system with integrated Active Air Spring

Figure 3.14 illustrates the improvement of the functional quality when replacing a passive component by an active one. Here, the benefit is an improved isolation function of a vehicle body from a country road's bumps. The picture results from a hardware-in-the-loop ride of a compact vehicle with a speed of 70 km/h along the road, cf. Fig. 3.14. The active component is the real hardware whereas the road and vehicle are virtual. For the passive system the standard deviation of the chassis acceleration is 0.81 m/s^2 . The asymptote is given for the ideal active system by 0.58 m/s^2 . This asymptote is formally derived by assuming an actuator with unlimited available actuating power and unlimited actuating speed. Figure 3.14 shows that, with the Active Air Spring, this level is sufficiently approached with a power consumption of roughly 10 W to 15 W for a compact vehicle such as the VW Golf. For reasons of dimension this power increases linearly with mass. The power consumption for a luxury vehicle having twice the body mass is 30 W for one vehicle corner and hence 120 W on average for the vehicle.

With regard to Figs. 3.5 and 3.14, the acceptability of the system arises by an increased quality of the functional performance. On the other hand, the energy and cost efforts are increased by the active components. Moreover, the additional components reduce availability due to their vulnerability as discussed above.

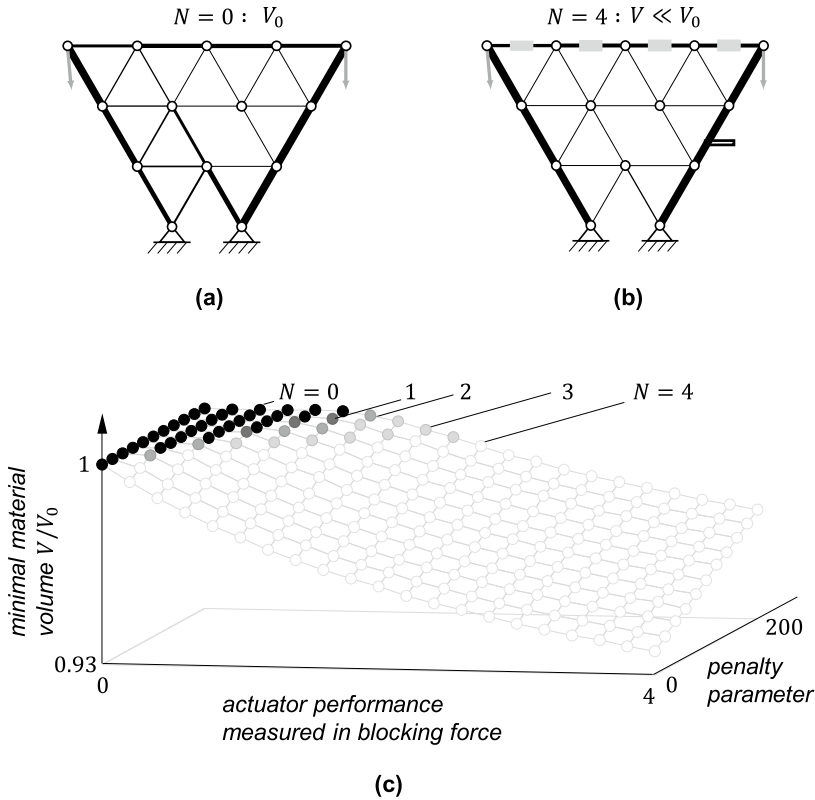


Fig. 3.15 Reduced effort measured in relative reduction of material consumption. **a** Robust optimised truss structure allowing only passive modules; **b** robust optimised truss structure allowing active components; **c** material saving with increasing number N of active modules and increasing actuator performance, cf. Sect. 6.1.2

Example—Truss structure with integrated actuator

In the above example, the active components serve to increase the functional quality of the system. From a different point of view, we may think of having the functionality unchanged and the active components serve to reduce the effort in material consumption.

This is exemplified by comparing two truss structures shown in Fig. 3.15a, b: Fig. 3.15a shows the optimised topology resulting from a robust optimisation of a passive, static system; Fig. 3.15b shows the structural optimisation for the same function, namely bearing defined load scenarios, when additional active components marked by a red box are integrated. Figure 3.15c shows the volume saving relative to the optimal passive system by replacing $N = 1, 2, 3, 4$ passive bars with active bars. The actuator performance is measured in the dimensionless actuating force ranging from 0 to 4.

3.5 Strategies for Mastering Uncertainty—Robustness, Flexibility, Resilience

Peter F. Pelz and Marc E. Pfetsch

In this section, we present three outstanding strategies or concepts for mastering uncertainty. These serve as a guide taking us on a tour throughout the book, i.e. we will introduce methods and technologies, cf. Chap. 5 that help us to implement the strategies. Chapter 6 is devoted exclusively to the three strategies.

The three strategies or concepts we identified to be most important in mastering uncertainty, cf. Chap. 6 are to

- (i) design and operate robustly,
- (ii) gain flexibility,
- (iii) enable resilience.

The so far cited demand for closing feedback loops may be seen as underlying to most aspects of the three strategies. This section serves to predefine, structure and exemplify the three strategies.

What distinguishes a robust system from a flexible and from a resilient one?

The need of the user is represented by the function g of the system. The function is always described by a verb such as ‘carry’ whereas the quality is given by the adverb, e.g. ‘carry safely’. I.e. the adverb represents the quality $F = \{F_1, F_2, F_3\}$ of how this function is fulfilled. We distinguish between a general quality and a functional quality. The general quality is to be understood as a generic term for minimum effort F_1 , maximum availability F_2 and maximum acceptability F_3 . These three sub-objectives result in the multi-criteria objective function. As explained in Chap. 1, acceptability can be achieved by the degree of function fulfilment, i.e. δg . This is the functional quality that is commonly understood as the quality of a product or system.

By concentrating on the function of the system and the effort required to achieve this function, we can now easily distinguish between robustness, flexibility and resilience. The differences are concisely collected in Fig. 3.16.

- (a) A robust system is characterised by the fact that the system fulfils one predefined function g with accepted functional quality δg , even if the function and resources have been described uncertainly or are uncertain themselves—as long as they are part of the specified uncertainty set—or even if the system is disturbed by uncertain external influences. For the constrained optimisation problem the objective function is given by $\min \{F_1, -F_2, -F_3\}$ where the effort is given by F_1 .
- (b) A flexible system is characterised by the fact that the system fulfils $i = 1, \dots, N$ predefined functions g_i with accepted functional quality δg_i . For the constrained optimisation problem the objective function is given by $\min \{\sum_i F_{1,i}, -\sum_i F_{2,i}, -\sum_i F_{3,i}\}$. Usually the effort F_1 is the most important task; i.e. several functions should be possible with minimal effort: $\min \sum_i F_{1,i}$.

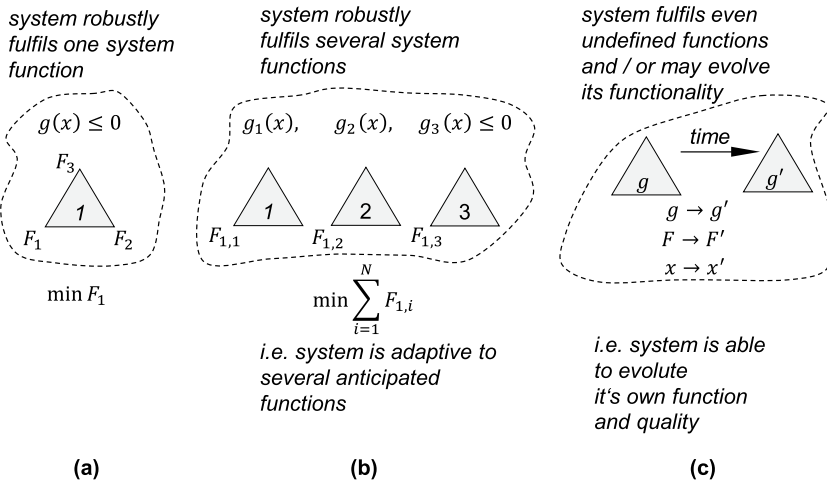


Fig. 3.16 Distinguishing robust, flexible and resilient systems from the perspective of the system’s functions: **a** a robust system does not only fulfil its function at the design point, but also in the surrounding neighbourhood, the so-called uncertainty set; **b** a flexible system does not only fulfil one function but several functions i.e. several design points; **c** a resilient system fulfils its function at the design point like a robust system but still enables a residual function when disturbed; the system has the inherent ability to recover

(c) A resilient system is characterised firstly by the fact that it does not only fulfil a predefined function g such as a robust system, but also retains a residual function g_{res} , if the system is disturbed at time $t = 0$. A resilient system may secondly show the ability to recover: from a distorted state, the system may recover in a time $t = 0, \dots, T$ to a function $g(x) \rightarrow g'(x')$ and/or a quality $F(x) \rightarrow F'(x')$. This function and quality may not be foreseen in the previous design and production phases. The design variants undergo transformation from x to x' during the evolution.

Robustness

Above we have defined the characteristics of a robust system. Here we would like to discuss how to establish a robust system. The comprehensive strategy is a methodology with many facets called robust design. Robust optimisation is a method needed in this methodology as we will demonstrate in a short example within this section and indeed in many examples throughout this book.

Robust Design goes back to Genichi Taguchi [14, 41]. In 1949, he started to develop the methodology of offline quality engineering, as Taguchi called it. The development of the method was accompanied by a project aimed at modernising the Japanese telecommunication system. The methodology proceeds from the system level to the product and process parameters. To reduce the influence of uncertain product and process parameters on the performance or functional parameters of

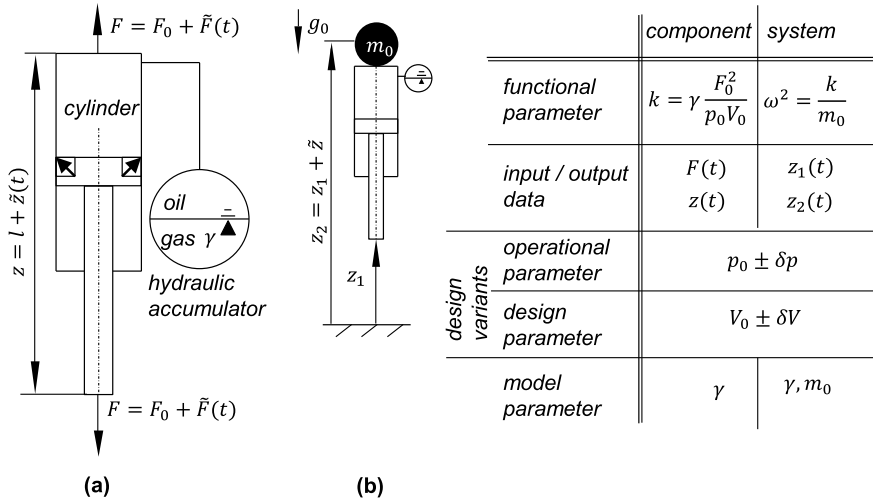


Fig. 3.17 Robust parameter design of a hydro-pneumatic suspension strut treated as a constrained optimisation problem; **a** shows the component and **b** the system; given is the list of functional parameters, input/output data, decision parameters and model parameters; note the change in parameters by broadening the view from the component to the system

the system, the sensitivity of design and process parameters are first determined by Design of Experiments (DoE). Following this first step called system design, the design variants are selected by robust optimisation. This second step is called parameter design. In the last step, called tolerance design, the tolerances for the design and operating parameters are defined.

Robust control in control theory is an approach to controller design. Hence, it has only one component of the system, the controller in focus. This controller shall provide its function for an uncertainty range of system parameters. Thus a robust controller is static in contrast to an adaptive controller, adapting itself to system's variations [1].

In conclusion, the robustness of a system is gained by a methodology. Per se it does not need any special effort measured in investment or material. Of course, it is important for engineers today to know and apply the methods of Robust Design. They are constantly being further developed. Thus, the robust optimisation described in this book lays the foundation for Robust Design.

3rd case study—Parameter engineering in the methodology of Robust Design

The following case study firstly illustrates the parameter design as understood by Taguchi. Secondly, the case study further illustrates the benefits of inherently robust design and operational concepts when discussing the various physical effects that can be used in a suspension system.

Figure 3.17 shows the hydraulic scheme of a hydro-pneumatic suspension strut known from mobile hydraulics, forming presses or from the chassis of the Citroen DS passenger car. The function of the system is given by three sub-functions $g = \{g_i, g_{ii}, g_{iii}\}$:

- (i) The first sub-function supports a chassis of mass m with the mass specific gravitational constant g_0 resulting in the time averaged force $F_0 = m_0 g_0$ at a distance l above the ground.
- (ii) The second sub-function, the levelling known from the Citroen DS, is to enable a constant distance l independent of the loading $F_0 = m_0 g_0$.
- (iii) The third sub-function is to isolate the chassis from vibrational excitation by the bumpy road.

This third sub-function is ensured by having a compliant system with the capability to store energy. Thus, the system has a natural frequency as discussed in Sect. 3.4. For an excitation frequency $\Omega^2/2 > \omega^2 := \frac{k}{m}$ the system is in isolation. Hence, the natural frequency ω_s is a predefined, i.e. specified value characterising the function of the system.

In Chap. 1 we discussed the objectives to be met when designing systems, cf. Fig. 1.9. All the three part-objectives, effort F_1 , availability F_2 and acceptability F_3 form the general quality of the system. The acceptability is fostered when the special quality of the system, i.e. the degree of the functional fulfilment is improved. Every physical-technical system that fulfils the functional description given above competes in quality. The latter is given by the variance, i.e. the stochastic uncertainty $(\delta\omega)^2$ of the square of the system's natural frequency. The better the desired natural frequency is hit by a technical system the better is the system's functional quality.

In a mass production we anticipate the uncertainty of the functional parameters, here the square of the natural frequency ω^2 as $\omega^2 = \bar{\omega}^2 \pm (\delta\omega)^2$. While the discrete variants are assumed to be already selected when defining the structure, the task of the parameter design introduced by Taguchi is to select the continuous design variants u in such a way that the constrained optimisation problem is solved:

$$\min_u (\delta\omega)^2 \text{ s.t. } \bar{\omega}^2 = \omega_s^2. \quad (3.5)$$

Here, ω_s is specified and thus fixed, whereas $\bar{\omega}$ and $\delta\omega$ both depend on the design variants u . The task function is clear: we would like to have a minimal uncertainty in the functional parameter, i.e. the eigenfrequency. In the competition of physical modules we compare three systems, coil spring, hydro-pneumatic suspension and air spring. A coil spring alone would allow the first and the last sub-functions but not the levelling function. Hence, the coil spring fails the competition. An air spring allows all the three sub-functions; and air springs are indeed standard for load-bearing systems when there is a large variance in the mass $m_0 \in [m_{\min}, m_{\max}]$ as it is given in commercial or rail vehicles. But air springs require a compressed air system including compressor, air dryer and filter to level the system. This results in a constant time averaged gas volume V_0 in a loaded condition independent of the loading F_0 and hence mass m_0 .

The needed effort, i.e. the investment in hydro-pneumatic systems is usually smaller. Here, the sum of gas and oil volume is constant within the suspension strut sketched in Fig. 3.17. With increased load F_0 oil is pumped by means of a pump into the system of constant volume. Hence, the gas within the hydraulic accumulator is compressed. The gas mass m_g does not change by the levelling.

For a minimal effort, we therefore would select in the first place a hydro-pneumatic suspension system. The design variants to be selected for the system are the absolute gas pressure p_0 of the accumulator and the gas volume V_0 of the accumulator in the unloaded condition, $u = (p_0, V_0)$. In our nomenclature p_0 is an operational variant for the manufacturing when the system is initially filled and V_0 is a design parameter. Both decision variants are uncertain. Presuming we know the incertitude in both decision variants, i.e. the operational variant of the production phase $p_0 = \bar{p}_0 \pm \delta p_0$ and the design variant $V_0 = \bar{V}_0 \pm \delta V_0$ we may solve the constrained optimisation problem.

Ignoring uncertainty of the use phase assumes a fixed suspended mass $m_0 = \text{const}$. The stiffness k of the suspension is the derivative of the force F compression z relation, i.e.

$$k := \left. \frac{dF}{dz} \right|_{F=F_0} = \gamma \frac{F_0^2}{p_0 V_0}. \quad (3.6)$$

Here γ is the isentropic exponent of the gas, i.e. a model parameter, cf. Fig. 3.17. With $k^2 = \bar{k}^2 + (\delta k)^2$ and the specified stiffness $k_s = \omega_s^2 m_0$ the constrained optimisation problem can be written as in an equivalent unconstrained problem by introducing the Lagrange parameter λ

$$\min_{p_0, V_0, \lambda} \left[(\delta k)^2 - \lambda(\bar{k} - k_s) \right] \quad (3.7)$$

(the two terms inside the square brackets sum up to the Lagrangian function). This problem is solved using Gaussian error propagation introduced in Sect. 3.2 for uncorrelated uncertainty

$$(\delta k)^2 = \left(\frac{\partial k}{\partial p_0} \right)_{(V_0, p_0)}^2 (\delta p_0)^2 + \left(\frac{\partial k}{\partial V_0} \right)_{(V_0, p_0)}^2 (\delta V_0)^2. \quad (3.8)$$

Hence, the parameter design, as Taguchi named it, yields the optimal value for the design variants:

$$\left(\frac{p_0}{\delta p} \right)_{\text{opt}} = \left(\frac{V_0}{\delta V} \right)_{\text{opt}} = \sqrt{\frac{\gamma}{k_s} \frac{F_0}{\delta p \delta V}}. \quad (3.9)$$

For this set of design variants, the product is said to be robust in the sense that the disturbances, i.e. the uncertainty in realising the pressure and volume, has the minimal impact on the uncertainty of the demanded functional property, i.e. the uncertainty of the stiffness.

So far Robust Design was applied and used on a component level; and as Fig. 3.17a indicates, it is the component for which the functional quality is indeed the stiffness. However, the quality the customer is interested in is not the quality of the component but the quality of the system. For the suspension system the functional quality is the natural frequency, Fig. 3.17b, and not the component's stiffness. Thus the sketched example is a typical example of how the system boundary, and hence the system view, cf. Fig. 3.3a, influence the design objective. For the—already disqualified—coil spring we have $\omega \propto 1/\sqrt{m_0}$. If we demand $(\delta\omega)^2$ to be small as a quality measure, storing energy by an elastic torsion of a beam—as it is done by storing energy in a coil spring—is not a physical effect ensuring robustness. The hydro-pneumatic suspension discussed above gives $\omega \propto \sqrt{m_0}$, i.e. the natural frequency increases with increasing load. This results directly from the given stiffness. Hence, both physical effects do now allow robustness with respect to mastering uncertainty in $m_0 \in [m_{\min}, m_{\max}]$.

There is a third principle to fulfil the three sub-functions. An air spring allows $\omega = \text{const}$ due to separating the sub-functions into (I) carrying and (III) isolating. In other words the air suspended system inherently uses three effects of functional separation. The load carrying sub-function is gained by the force balance $F_0 = (p_0 - p_a)A = m_0 g_0$ with the ambient pressure p_a and the springs cross-sectional area A . The levelling is done by adapting the pressure p_0 to the loading. The air spring's stiffness is

$$k := \left. \frac{dF}{dz} \right|_{F=F_0} = \gamma p_0 \frac{A^2}{V_0}. \quad (3.10)$$

Hence, the natural frequency of an air spring with $p_0 \gg p_a$ is

$$\omega^2 = \frac{k}{m_0} \approx \frac{1}{\gamma} \frac{g_0}{\frac{V_0}{A}}. \quad (3.11)$$

It is independent of the uncertain model parameter $m_0 \in [m_{\min}, m_{\max}]$ and the system is robust with respect to uncertainty in the usage phase at least in this one aspect. Typically, the height of a cylindrical spring $\frac{V_0}{A} \approx 350$ mm. This results in a natural frequency of 1 Hz. Indeed, if we compare springs in different applications, such as chassis of trains, commercial vehicles or passenger cars, they all show the same height of roughly 350 mm. The height is adapted to the sensitivity of adult humans regarding vibrations. The design is inherently robust with respect to the uncertainty in the usage phase.

Employing inherent robust design and operating concepts

We have learned from the above example that sometimes uncertainty can be overcome by choosing an inherently robust design and/or an inherently robust operating concept. In this example, energy storage in a gas volume is chosen as the physical effect. In the selected operating concept, the time-averaged gas volume (not the gas mass) is constant. This means that all three sub-functions mentioned above can be fulfilled independently of each other: carrying, levelling and isolating. In summary, the separation of functions was chosen as a concept to achieve robustness.

Separation of function is not the only concept inherently leading to robustness. Here we give a short, certainly incomplete list of design principles that are familiar to many engineers. Each concept fosters robustness and is hence a way how to master uncertainty. The seven design principles that inherently lead to robust construction solutions are to:

- (i) enable overload protection,
- (ii) enable overload capacity,
- (iii) enable self-adaptation to increasing loads,
- (iv) enable compensation of uncertainty,
- (v) enable self-healing,
- (vi) separate functions and
- (vii) close feedback loops.

Each design concept and each concept of operation are illustrated by some short examples:

First (i), the *overload protection* is provided, for example, by a pressure relief valve or a split pin with a defined shear force. Cellular solids or foams [2] show an overload protection provided by the material itself. As long as the foam still has a compression margin, the compression force is basically determined by the buckling load of the cell walls. The foamed elastomer sole of jogging shoes enables—as desired—high energy absorption with a small force amplitude. The same principle, high energy adsorption with small force amplitude in the force-displacement diagram, is used for crash structures and railway buffers.

Second (ii), the *overload capacity* is made possible by additional energy storage. As to plastic materials, the energy storage is the irreversible deformation work with overload. As is well known, brittle materials tend to fail spontaneously. This principle applies in general. Supply and energy chains become robust if storage devices are integrated.

Third (iii), the *self-adaptation* to increasing loads is made possible e.g. by an O-ring: the higher the fluid pressure the higher the sealing contact pressure between the elastomer and the solid surface. The elasticity is a prerequisite for this elegant way of mastering uncertainty in the usage phase. This prerequisite was lost in the Challenger space shuttle disaster [52], because the elastomer of an O-ring tank seal had been frozen during a cold night. In the glass state, the material is frozen in a deformation state without the necessary contact stress. The leakage of fuel caused the explosion. The Challenger disaster is one more example of model uncertainty

ignoring either the influence of the temperature on the function of the seal or ignoring the temperature as a model parameter.

The design concepts of overload protection, overload capacity and self-adaptation, especially as self-enhancement and self-repairing principles, are also used as a basis for the realisation of resilient system properties, cf. Sect. 6.3.2.

Forth (iv), the *compensation* of uncertainty is reached by integrating elastic elements at the interfaces of system components. Foil air bearings are typical for this. The integration of the bumpy elastic foil compensates uncertainty in the shaft and the journal diameter.

Fifth (v), *self-healing* materials or structures are the dream of mechanical engineers. Unfortunately these are very rare. A current example is a modern mountain bike tire. The air tube is replaced by a liquid sealant which is added to the inside of the tire. Every small puncture is self-healed by this sealant.

Sixth (vi), the *separation of functions* is the most basic concept to master uncertainty. A clear assignment of functions to components includes, for example, the avoidance of double fitting. The two functions sealing and load-bearing are known to be separate. A piston seal should not have a load-bearing function and a guide ring should not have a sealing function. In the above mentioned examples it is interesting to note that the separation of functions is a general concept to gain robustness of a system not only in the design and operation of technical systems: all modern forms of government use the principle of separation of powers [9]. Modern business organisations also identify roles, i.e. functions to gain robustness in many aspects including legal ones.

When designing load-bearing systems, the concept of separation of functions is applied to the flow of forces and is referred to as the design for clarity, cf. Sect. 6.1.6. The clarity of force flow significantly supports the mastering of accumulated uncertainty from the product life cycle processes.

Seventh (vii) *closing the feedback loops* is always a good advice: not always all facets of mastering uncertainty have to be considered: in nature, feedback control is a very successful concept to manage uncertainty; in Chap. 1 we pointed out the benefit to close feedback loops over the phases of the product life cycle. For feedback control, the model uncertainty discussed in Sect. 4.3 is only of minor importance. In fact, feedback control works even if there is no model of the system. On the other side, data uncertainty, i.e. the uncertainty of the measured or calculated data to be fed back, is most important when closing control loops. The situation is somewhat inverse in forward control. Here, in fact the model uncertainty and model quality are most important.

The given list of inherent robust design and robust operating concepts is not complete. The presented concepts are subliminally present in every design task. It is inherent to the optimisation and learning, not only in mechanical engineering.

Illustrating examples are e.g. the frequent weekly cooking of spaghetti and the rare roasting of a piece of meat. The latter may happen only for Christmas once a year. Only if the signal-to-noise-ratio when tasting salt is high enough, feedback will improve the cooking result of spaghetti from time to time. A typical feedforward control exists when the roast time of the roast has to be estimated in advance. This requires a model of the oven, temperature control and a model of the roast. The

model of the roasting is parameterised by the weight of the meat piece. A good cook has such a model in mind. Again the model does not need to be in a mathematical form even though it might be so. If the model uncertainty is too large, the result will be of poor quality. The open control loop can be closed by a roasting thermometer, so that uncertainty is controlled without using a model at all.

Robustness, just like the mastering of uncertain loads, also requires insensitivity to disturbance parameters that affect the system externally. Knowledge of the influences on the system from its environment and vice versa is an essential prerequisite for development. The detailed analysis of disturbance influences by means of a process model is described in Sect. 5.2.3. In Sect. 5.1.2 it is shown how the documentation of desired functional properties and expected disturbance parameters positively influences the development process. The choice of physical effects that are insensitive to the expected disturbances can fundamentally contribute to the robustness in the early phases of the system design as shown in Sect. 6.1.5.

Flexibility

A flexible system is characterised by the fact that the system fulfils $i = 1, \dots, N$ predefined functions g_i with accepted functional quality δg_i .

Recapturing the above given characterisation, the question arises how this flexibility can be achieved. There are several ways how to gain flexibility and there are many perspectives on the topic, cf. the review article about gaining flexibility in manufacturing [46]. As before, our perspective is a more general one being suitable to all phases of the product life cycle and also, as before, we have the systematic engineering design perspective [36]. From the functional view, we either can separate functions in different physical or software modules, or we can integrate functions in one physical module which may contain software modules.

Smart modularisation or smart modules, both concepts may enable flexibility

In the above explanations, we present the separation of functions as an inherent robust design and/or inherent robust operating concept. This applies at the level of a functional unit or a single component. Recognising that a module integrates some functional units, we have the opportunity to discuss the degree of functional integration in the module. Today the term “smart” module is ubiquitous. Therefore we should discuss the degree of “smartness”. This degree is determined by the degree of integration or separation of functions.

In the following, we first differentiate between the two concepts to gain flexibility (i) modularisation by the smart separation of functions and (ii) smart modules by integration of function, see Fig. 3.18. This is followed by a more detailed look at the two contrasting concepts on how to gain flexibility.

Figure 3.18a shows flexibility gained by separation of functions into m sub-functions and the associated m physical or software modules. The matrix gives the utilisation n_j of the different modules. The effort F_1 is minimal if the total utilisation $\sum_j n_j$ is maximised and at the same time the number of modules m is minimised:

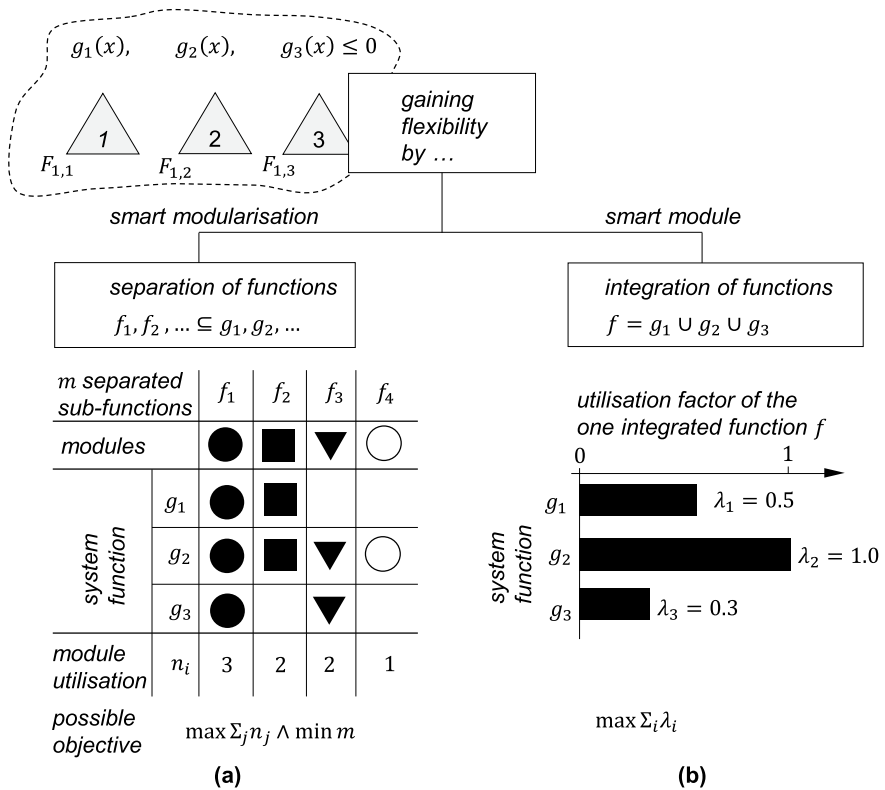


Fig. 3.18 **a** Flexibility gained by smart separation of functions into m sub-functions and the associated m physical or software modules; **b** flexibility gained by integration of function of the g_1, g_2, \dots user-defined individual functions into on master function realised by one smart module

$\max \sum_j n_j \wedge \min m$ subject to that in practise physical modules count more than software modules. The number of modules m is a measure of the internal complexity with regard to the design and the production phase. The external complexity comes from the user’s view, e.g. the user would like to have common parts for maintenance. At the same time the specific user would like to have his or her needs matched fully by the system. Hence, the user experience should be kept in mind when designing a modular product.

Figure 3.18b shows how flexibility is gained by integration of function of N user defined individual functions into one master function realised by one “smart” module. For the average user the system shows too much functionality, which is of course not bad for the user but may be costly for the producer. The part of the functionality used by one user is denoted by $\lambda_i \leq 1$. The objective for the integration of function is hence $\max \sum_i \lambda_i$.

So far it has become clear that the two asymptotes of flexibility are first (i), “smart” modularisation and second (ii), “smart” modules. The first asymptote is a methodology, the second one is a technology, cf. Chap. 5.

Smart modularisation

We discuss modularisation as a smart separation and integration of functions into standardised modules as a strategy of flexibility and hence of mastering uncertainty. A module is a physical unit with several integrated functional units. The module can be used to fulfil different system functions by combining it with other modules. Thus, a component is formed out of several standardised segments. This concept of combining standardised modules was already used in the 13th century when standardised stone modules were prefabricated in a quarry to form structures of gothic cathedrals in medieval France. Today it is the basic principle to gain flexibility for example in ship building, process engineering and many other fields of engineering.

Modularisation is fostered with standardised interfaces for energy, forces, displacements and information; e.g. in the *Austauschbau*—being common in mass production, Chap. 1,—the uncertainty in geometric dimension is standardised by means of tolerance fields. Obviously by modularisation and standardisation the division of labour is fostered and the separation into the single elements of a supply chain is enabled. The development speed may be equally increased by integrating known modules into the system. The uncertainty of those modules is usually quantified. In some cases it is even possible to gain a legal conformity certification on the module level. If the module uncertainty is mastered together with the uncertainty on the basis of the interfaces, the system's functionality and the objective functions can be quantified.

It is self-evident that sources of uncertainty arise in the modularisation, in the first instance there are modules, or in the second instance there may be interfaces. Based on our experience the availability of a technical system is often limited by the failure of electrical power or signal connectors. Hence, not only the standardisation of the interfaces but also their robustness are most important when designing modular systems. The robustness of the interface is fostered by obeying the above five inherent robust design and operation concepts. So the concept (iv), compensation of uncertainty, may be realised by a flexible interface. As a rule, the interface, or more generally speaking, the module's boundaries should be selected in such a way that a module's sub-function and quality are only marginally influenced by a detailed location of the boundary. This rule is exemplified in the following case study.

4th case study—Smart modularisation of a control valve

In this case study we address the question of how to solve the constrained optimisation problem formulated in Fig. 1.9b: with minimised effort subject to the function g_i being fulfilled.

$$\min_x \sum_{i=1}^N F_{1,i}(x) \quad \text{s.t.} \quad g_i(x) \leq 0, \quad i = 1, \dots, N. \quad (3.12)$$

For this purpose, a product has to be evaluated first from the different stakeholder views. At least there is (a) the manufacturer and (b) the customer. In a business to business market there is often also (c) the planner working as a service provider and (d) the approval authority. All two or four stakeholders have different interests. Hence, the optimisation problem is more complex:

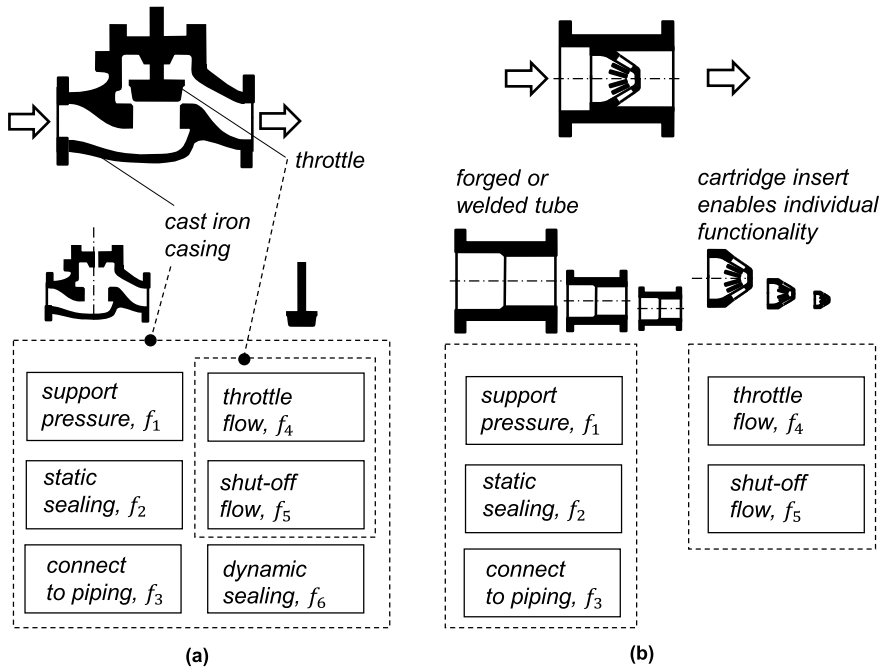


Fig. 3.19 Control valve design: **a** the integration of functions in a control valve hinders modularisation and increases internal and external complexity. Internal complexity may be measured by the number of cast moulds needed to meet the required functions. External complexity is measured by the different spare parts needed for maintenance. **b** Separation of functions enables a size ranged modular design of the valve reducing internal and external complexity

$$\min_x \sum_{i=1}^N \left[w_a F_{1,i}^a(x) + w_b F_{1,i}^b(x) + w_c F_{1,i}^c(x) + w_d F_{1,i}^d(x) \right] \text{ s.t. } g_i(x) \leq 0, \quad i = 1, \dots, N. \quad (3.13)$$

Hereby, it becomes clear that the result of a modularisation strategy depends on the weights w_a , w_b , w_c , w_d of the individual stakeholder.

Figure 3.19 gives an example of a modularisation derived from a functional integrated control valve typically used in the process, energy or petrochemical industry. The functional separation as presented in Fig. 3.19b is obviously the prerequisite for the independent size ranged modular design of the forged or welded tube and the cartridge insert. The functional quality is determined by the insert allowing a detailed design to fulfil the primary prescribed function of the valve, i.e. throttle the flow. The tube segment fulfils the secondary functions only and this is made possible in a wide range independent from the insert. The interface between the module “tube” and the module “cartridge” allows large tolerances being important for successful modular designs, i.e. the interface is robust in the above prescribed sense. Looking at the tube, it is much simpler in design and manufacturing compared to the cast iron casing of the traditional design. This reduces the uncertainty in the supply chain and shortens delivery times. Overall the customer experience with the entire numerous facets can be improved by smart modularisation.

Mastering module uncertainty by smart test and development methods

The case study in Sect. 3.3, the ‘energy ship’, was motivated by the fact that combining modules with known uncertainty to build a new system enables mastering uncertainty. In fact that is what we observe today in the automotive industry where variants are derived from platforms and the platforms are based on modules. Hence, it is worthwhile shedding some light on smart test and development methods when talking about smart modularisation.

We focus on the following concepts to master the module uncertainty:

- (i) early failure on the basis of a model test,
- (ii) scaling of uncertainty in a size ranged modular design,
- (iii) module in the loop.

The first concept is a very easy one being part of the agile development. Despite the overall simulation methods we have today, we should start very early in the development phase to have virtual and real mock ups but also – being even more important—physical functional models that allow the evaluation of the expected functions, functional quality and general quality. The second concept is based on the similarity principle to derive uncertainty of a scaled prototype from the model test, cf. Sect. 4.3.6. The third concept is widely used today. The uncertainty in function and quality of a module can be evaluated by integrating a real module into a virtual model of the system.

Figure 3.20 shows this concept known as hardware-in-the-loop (HiL). Instead of integrating a new or adapted module into a real system all at once, Fig. 3.20(i), the module is encapsulated by an active physical interface to the cyber world, thus simulating the overall system, Fig. 3.20(ii). This strategy evaluates the functional quality of new modular components with reduced effort and improved testing possibilities. The difficulty here is on the one hand the model uncertainty of the overall system and on the other hand the uncertainty in the flow of mass, energy and information between the test module and the active interface.

Smart modules

Instead of following the concept of smart modularisation we may think of satisfying a “power user” or a “power application” with only one smart module. Hypothetically, all the functions g_1, g_2, \dots shown in Fig. 3.18b could be represented in one system or functionally integrated module. Thus, the module is oversized for most applications; but in the overall picture this strategy can be advantageous with regard to the total effort $\sum_i F_{1,i}$. Such smart modules were already discussed in the past in the context of flexibility in manufacturing [46] but only now, with the help of control theory, these smart modules become reality: The ability of the smart module to adapt to a specific application g_i is necessary. This adaptivity often requires an actively controlled process. In the context of this book, the Active Air Spring presented in Sect. 3.6.2 or the 3D Servo Press presented in Sect. 3.6.3 are examples of smart modules. Since the integration of functions into smart modules will accompany us throughout the book, we have decided to keep this subsection short.

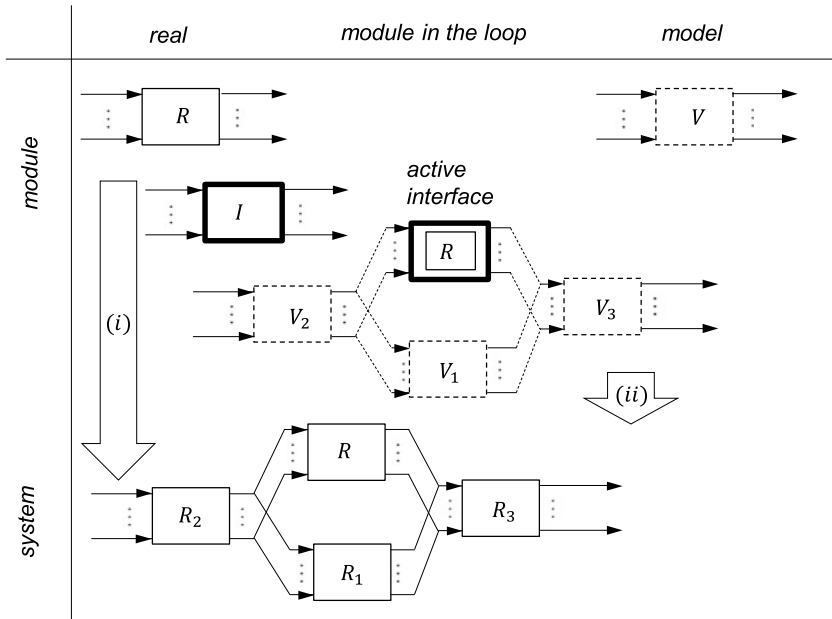


Fig. 3.20 Mastering uncertainty by stepwise integrating a module into a real system in combining the cyber world, with virtual modules V , and the real world, with real modules R by active interfaces. The interface I itself is an active component in the above defined sense. This concept is known as hardware- or module-in-the-loop (HiL, MiL)

Resilience

Again we recapitulate the characterisation of a resilient system, cf. Sect. 6.3 as described above:

A resilient system is characterised firstly by the fact that it not only fulfils a predefined function g like a robust system, but also retains a residual function g_{res} , if the system is disturbed at time $t = 0$. A resilient system may secondly show the ability of recovery: from a distorted state, the system may recover in a time $t = 0, \dots, T$ to a function $g(x) \rightarrow g'(x')$ and/or quality $F(x) \rightarrow F'(x')$. This function and quality may not be foreseen in the previous design and production phases. The design variants transform during the evolution from x to x' .

In contrast, robust and flexible systems allow either the fulfilment of one function or several system functions. In both cases, i.e. for robustness and flexibility, the system's functions may be uncertain. Sometimes it occurs that the function, i.e. the usage phase, is in part unknown when designing and manufacturing the system. Sometimes a system only fulfils a partial function due to an accident or a catastrophe.

Seeing resilience as strategy to master ignorance or nescience and flexibility or robustness as a method to master incertitude or stochastic uncertainty is one possibility. But this classification shows a difficulty, as the result of the classification depends on the degree of information on the usage scenarios which allow us to distinguish

between resilience and the other strategies. Hence, following this classification the difference between robustness and resilience is fuzzy at least in a strict scientific context: the transition between robustness and resilience is smooth and the use is determined by the context.

However, there is also another, namely a sharper distinction between the concepts of robustness and resilience. More precise is the characterisation from the perspectives of the system function and the system quality measured in effort, availability and acceptability: a resilient system is one that still shows a residual function when heavily disturbed. On top of this, a resilient system may have the ability to evolve its own function and quality. As pointed out in this chapter, evolution is a time-demanding process.

In fact, often this time-demanding process is commonly associated with resilience. Consistent with this understanding, we rather ask for the sub-functions or characteristics a system shall have to allow for the evolution of its functionality and quality from an initial function g , quality F associated to the design and operation parameters x at a time t to g' , F' , x' at a time t' : $x \rightarrow x'$, $g(x) \rightarrow g'(x')$, $F(x) \rightarrow F'(x')$.

This evolution requires sub-functions that are called resilient in a system. It should be able to

- (i) measure,
- (ii) react,
- (iii) learn,
- (iv) anticipate.

The four functions can be represented by human or “smart” process chains. Depending on the process chain and system boundaries, man and machine can work together as partners or independent of each other.

There are some sequences in our mind set: For the evolution mentioned, flexibility may be a further prerequisite. Hence, we do have the understanding of resilience \supset flexibility \supset robustness. If we evaluate the four required sub-functions, there is again a sequence with (a) measure is easier than (b) react, is easier than (c) learn, is easier than (d) anticipate.

Chapter 6 is devoted to the in-depth introduction of the three strategies.

3.6 Exemplary Technical System Mastering Uncertainty

Maximilian Schaeffner

The theoretical considerations should equally be practically valid, verifiable and falsifiable. This applies especially to those methods and strategies outlined in the previous sections for our approach to master uncertainty in load-bearing structures of mechanical engineering. In particular, we have introduced methods and strategies for the analysing, quantifying and evaluating as well as finally mastering uncertainty along the product life cycle from the design via production to usage. In order to investigate and prove the effectiveness of our approach, it is crucial to apply the methods and technologies to exemplary technical systems. Therefore, technology

demonstrators were designed and investigated both numerically and experimentally. More details are disclosed in the upcoming chapters.

Three of these technology demonstrators are introduced in greater detail in the following subsections. Section 3.6.1 provides an introduction of the Modular Active Spring-Damper System as a generic load-bearing system similar to an aircraft landing gear that captures a wide variety of applications. In Sect. 3.6.2 we describe the Active Air Spring as an active module to enable disturbance compensation in an automobile suspension. Finally, in Sect. 3.6.3 we introduce the 3D Servo Press, which is a new press concept to increase flexibility and productivity in forming processes.

The three technology demonstrators shown in Sects. 3.6.1–3.6.3 stand as *pars pro toto* for all technologies presented within this book. They render the application of the methods and strategies possible to analyse, quantify, evaluate and finally master uncertainty across all phases of the life cycle. Examples are the identifying of data-induced conflicts in Sect. 4.2, quantifying model uncertainty in Sect. 4.3 and evaluating resilience in Sect. 6.3.

3.6.1 Modular Active Spring-Damper System

Christopher M. Gehb, Maximilian Schaeffner, Robert Feldmann,
Jonathan Lenz, and Tobias Melz

The Modular Active Spring-Damper System (German acronym: MAFDS—Modulares aktives Feder-Dämpfer System) is a generic load-bearing system that serves as a platform for applying and testing methods and technologies to master uncertainty, compare Chap. 2. The motivating origin is based on an aircraft landing gear which reflects in particular the conflict between maintaining the main dynamic features, i.e. load bearing, load distribution, structural stabilisation and vibration control, combined with lightweight design. The investigations to describe, evaluate and master uncertainty are derived on a virtual MAFDS using phenomenological and mathematical models as well as on the physically realised structure of the MAFDS, see Fig. 3.21.

The MAFDS illustrates the possibilities of mastering uncertainty in a realistic and descriptive manner for scientific purposes. The mechanical requirements of the MAFDS are largely, but not exclusively, based on those of an aircraft landing gear and may also capture the quarter car dynamic behaviour of an automotive chassis. Nevertheless, the MAFDS and its test environment do not intent to replace, supplement or extend the existing industrial and product-oriented test procedure of commercial aircraft landing gears and automotive chassis or the aircraft landing gears and automotive chassis itself. In fact, the findings from applying and testing methods and technologies to master uncertainty in load-bearing systems are supposed to be transferable on scientific scales to many other load-bearing systems. Said findings serve to increase the acceptance and credibility of the investigated methods and tech-

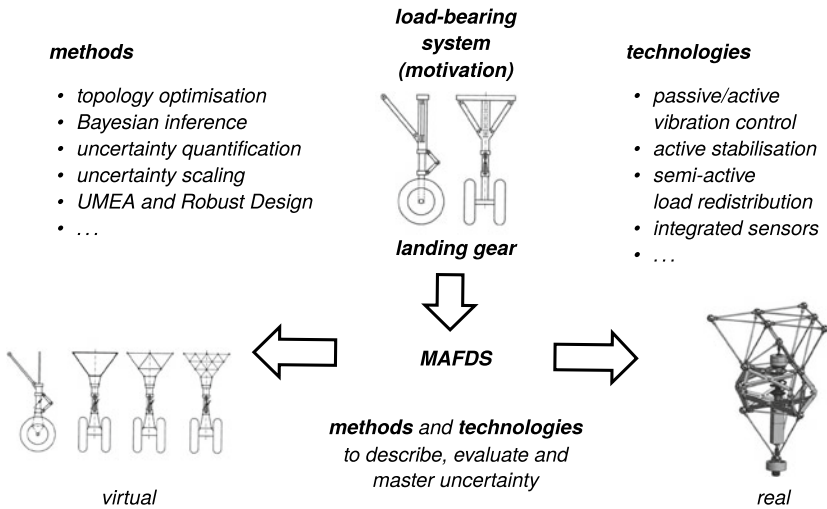


Fig. 3.21 Methods and technologies for mastering uncertainty in the generic load-bearing system MAFDS

nologies to master uncertainty. This uncertainty to be mastered includes all forms of uncertainty introduced in Chap. 2.

The virtual part comprises the methods and technologies to describe, quantify and evaluate the uncertainty, which means to examine the phenomenological and mathematical models and simulations to finally master the uncertainty [19]. These include e.g. investigation of topology variations in load-bearing structures that cannot be experimentally implemented due to the large number of possible variations. Furthermore, this part involves to apply and test methods, such as Robust Design, Uncertainty Mode and Effect Analysis (UMEA) in Sect. 5.2.1, process and uncertainty modelling, and model parameter calibration as in Chap. 4.

The physically realised part comprises components that are constructed, fully realised and experimentally tested, such as the truss structures, supports, beams and spring-damper, compare Fig. 3.22. Due to its modular structure, the virtual as well as the physically realised MAFDS enable the integration of different technologies to master uncertainty through e.g. load redistribution, stabilisation and vibration control. Therefore, various passive, sensory, semi-active and active components are applicable for each intended process manipulation. In the following, the possibilities of sensory, semi-active and active process manipulations for the MAFDS are presented after the test setup and the MAFDS are introduced in detail.

Test setup of MAFDS

The MAFDS is a load bearing system that can be used as a passive system and allows for the integration of semi-active and active components due to its modular setup [11]. The passive structure and its components are shown in Fig. 3.22a. It

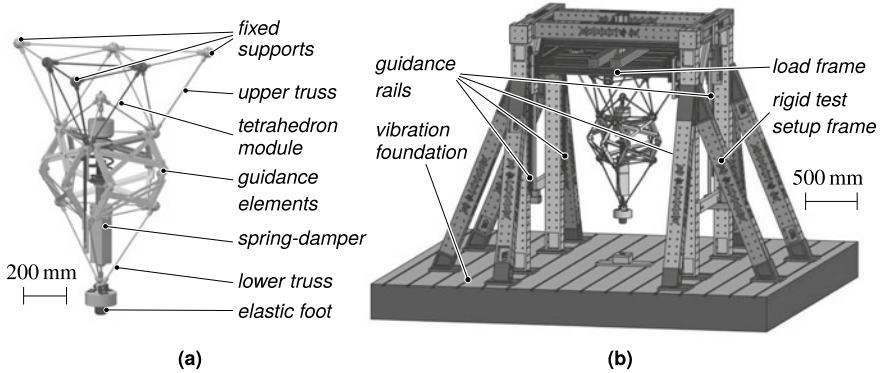


Fig. 3.22 **a** MAFDS, **b** test setup with MAFDS; not shown is the installation of the MAFDS in a servo-hydraulic test rig

consists of an upper truss structure, a lower truss structure with an elastic foot, guidance elements and a spring-damper. The three supports of the upper truss are fixed to a load frame that guides the MAFDS in vertical direction for experimental drop tests, see Fig. 3.22b. Additionally, it is possible to install the MAFDS in a servo-hydraulic test rig in order to apply base excitation, see Fig. 3.29 upper right picture. By this means, we realise a defined frequency excitation and e.g. road excitation. However, this section focuses on the drop test setup.

The upper truss consists of four tetrahedron modules with slender beams and solid nodes, of whom one is highlighted in Fig. 3.22a. The individual beams can be easily exchanged by sensory, semi-active or active beams, as presented in Sects. 5.3.6, 5.4.6 and 5.4.7. The four tetrahedron modules are linked to each other, yielding the upper truss structure. The lower truss of the MAFDS consists of one tetrahedron module. An elastic foot and an additional mass are attached to the lower node to represent the stiffness and the inertia of a wheel of an air plane landing gear or a car.

The dynamic behaviour of the MAFDS is governed by the properties of the spring-damper. For the application in the MAFDS, the passive suspension strut of a mid-range car serves as a spring-damper that is connected to the upper and the lower truss via a torque-free connection.

In order to enable the relative translation of the upper and lower truss structures in vertical direction, both are connected by three kinematic guidance elements. The design and position of the three guidance elements allow for simultaneous transmission of the transverse forces and bending moments between the lower and upper truss. In this way, the spring-damper is not charged with transverse forces or bending moments. Thus, the absorption and dissipation of impact energy and the low-frequency vibration reduction mostly take place via the spring-damper. A simple two degree freedom (2 DOF) model of the MAFDS to capture its dynamic behaviour will be presented in Sect. 4.3.3.

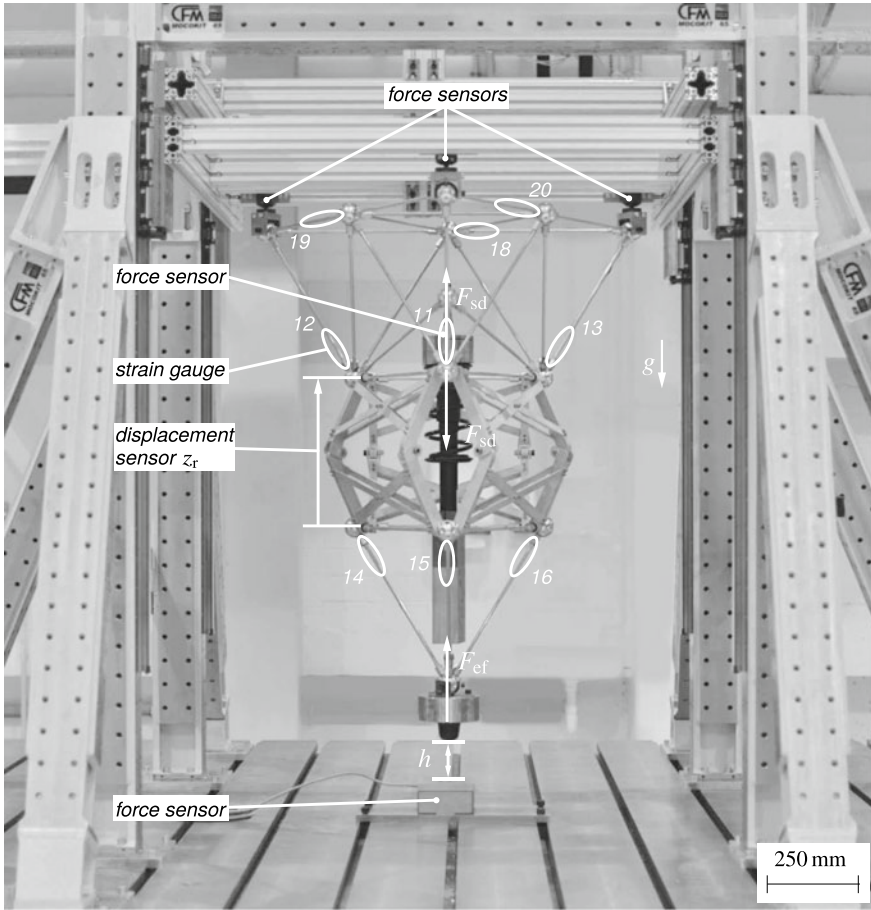


Fig. 3.23 MAFDS and its sensors and selected measured quantities [13]

The test setup used for the experimental testing of the MAFDS is depicted in Fig. 3.22b. It consists of a rigid test setup frame attached to a vibration foundation. Parallel guidance rails are mounted on the test setup frame and enable a low-friction vertical movement of the load frame that can translate along the guidance rails. The test setup enables the introduction of static as well as dynamic loads, the latter via drop tests of the MAFDS. These are carried out in a similar way as to landing gear tests. After the load frame with the MAFDS attached to it is lifted up to a desired drop height h as depicted in Fig. 3.23, the load frame is released for a drop test. Additional weights m_{add} can be added to the load frame, similar to varying loads of an airplane or a vehicle. The deterministic, but varying input quantities shall induce data uncertainty into the system, as defined in Sect. 2.1. The drop height h and the added mass m_{add} thus constitute the inputs to the experimental drop tests.

Table 3.1 Input combinations of added mass m_{add} and drop height h as well as resulting number of drop tests, upright and tilted configuration possible

m_{add} in kg	h in mm	Number of drop tests
0	10, 20, 30, 40, 50, 60, 70, 80, 90	9×5
10	10, 20, 30, 40, 50, 60, 70	7×5
20	10, 20, 30, 40, 50, 60	6×5
40	10, 20, 30, 40, 50	5×5
60	10, 20, 30, 40	4×5
80	10, 20, 30	3×5
100	10, 20	2×5
		$N = 180$

Additionally, MAFDS can be tilted by an angle in direction of each of the three fixed supports of the upper truss. Thus, it is possible to introduce lateral forces into the elastic foot. For the studies presented in this book, $N = 180$ experimental drop tests were carried out and measurements taken accordingly. In Table 3.1, the input combinations and respective repetition of measurements are shown. The drop tests specified in Table 3.1 were conducted for the MAFDS in an upright position as well as tilted in three different directions.

In order to capture the dynamic behaviour during the drop tests, the MAFDS is equipped with a comprehensive set of sensors to measure forces, bending moments, strains and displacements. Selected sensor positions are shown in Fig. 3.23. The forces at the three fixed supports, Fig. 3.22a, where the upper truss of MAFDS is connected to the load frame, are measured via triaxial piezoelectric force sensors. Normal and bending strains are measured using strain gauges attached to selected beams of the upper and lower truss. The locations of the strain gauges are indicated with ellipses in Fig. 3.23. The force between the spring damper and the upper truss is measured by a uniaxial strain-based force sensor and is denoted by the spring-damper force F_{sd} . The impact force that the elastic foot of the MAFDS exercises on the vibration foundation during the drop tests is measured by a triaxial strain-based force sensor with its vertical component being denoted by the elastic foot force F_{ef} . Furthermore, the relative displacement between the upper and the lower truss is measured using displacement sensors being denoted by z_r . Exemplarily, Fig. 3.24 shows measurements of the relative compression z_r , the spring-damper force F_{sd} and the elastic foot force F_{ef} for a drop test with zero additional weight $m_{\text{add}} = 0$ kg and drop height $h = 0.09$ m, where the peak values $z_{r,\text{peak}}$, $F_{\text{ef,peak}}$ and $F_{\text{sd,peak}}$ are marked.

Using the measurements of the MAFDS, the dynamic behaviour and aspects related to data uncertainty and model uncertainty (see Sects. 2.1 and 2.2.2) are further investigated in Sects. 4.2.3 and 4.3.3.

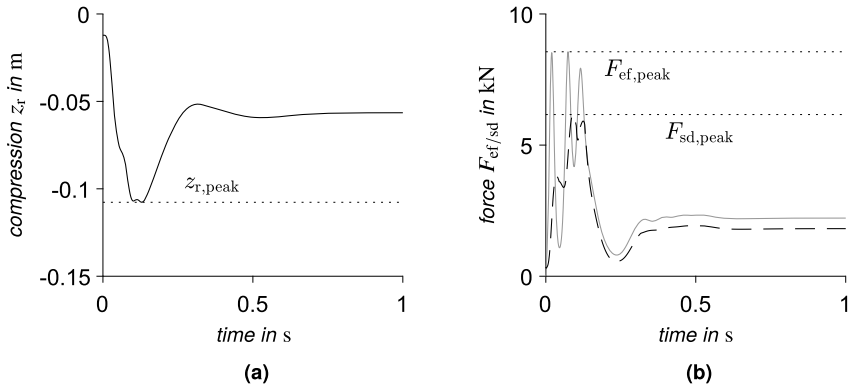


Fig. 3.24 Outputs from measurements of a drop test: **a** relative compression z_r (—), **b** force in the elastic foot F_{ef} (—) and force on the spring damper F_{sd} (— —)

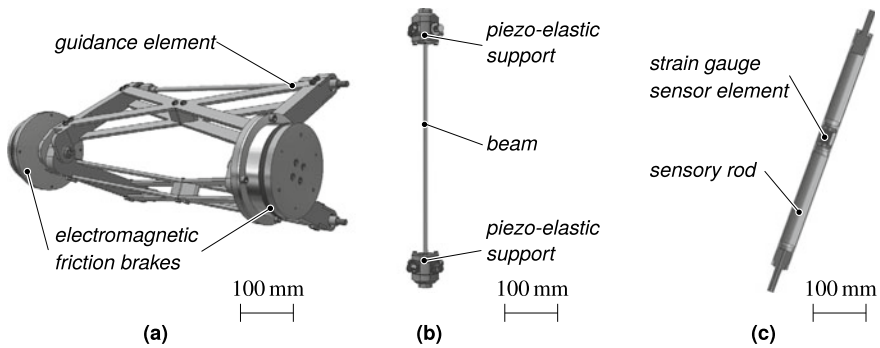


Fig. 3.25 Technologies for mastering uncertainty in the MAFDS by semi-active and active process manipulation: **a** semi-active guidance elements for load redistribution, Sect. 5.4.8, **b** beam with piezo-elastic supports for active buckling control, Sect. 5.4.7, and (semi)-active piezoelectric shunt damping, Sect. 5.4.6, **c** sensory rod for condition monitoring, Sect. 5.3.6; not shown is the Active Air Spring presented in Sect. 3.6.2

Semi-active and active process manipulation in the MAFDS

So far we have presented the design of the passive MAFDS that is used to quantify and evaluate uncertainty in a generic load-bearing system. In addition, the modular structure of the MAFDS allows for the possibility to modularly exchange sensory, semi-active and active components in order to master uncertainty within the usage phase, see Sect. 3.1. The technologies were developed as individual systems and validated in component tests, as described in detail in Sect. 5.4. Exemplarily, Fig. 3.25 shows three different technologies that can be integrated into the MAFDS to master different sources of uncertainty being present during the usage phase of the MAFDS.

In a new concept for load redistribution, compare Sect. 5.4.8, the semi-active kinematic guidance elements presented in Fig. 3.25a are used to influence the load

path of the MAFDS, e.g. in a drop test. Thus, the guidance elements between the upper and lower truss structures of the MAFDS are enhanced with a new dynamic function beyond the original kinematic one. Innovative piezo-elastic beam supports for beams with circular cross-section are shown in Fig. 3.25b. In one application, the piezo-elastic supports are used for active buckling control of axially loaded beam-columns in the upper truss of the MAFDS to increase the load-bearing capacity of the upper truss structure, Sect. 5.4.7. In another application, they are used for the attenuation of lateral beam vibrations within the truss structures of the MAFDS by piezoelectric shunt damping, Sect. 5.4.6. Figure 3.25c depicts a sensory rod with integrated strain gauge sensors to monitor the load status of individual beams in the truss structure of the MAFDS, which is manufactured using incremental forming processes, Sect. 5.3.6.

Further technologies, which can be integrated into the MAFDS, but are not shown here, are the passive vibration control by a spring-damper with integrated hydraulic vibration absorber; this is an alternative to the passive spring-damper of the MAFDS, Sect. 5.4.4, and further the active vibration control of the MAFDS by the Active Air Spring, which is presented in detail in Sects. 3.6.2 and 5.4.5.

Thus, the modular structure of the MAFDS allows the numerical and experimental testing of passive, sensory, semi-active and active technological measures to master the uncertainty during the usage phase of the MAFDS. These technologies are integrated into the MAFDS and tested in different load scenarios analogue to the passive MAFDS, see Table 3.1. By comparing the passive structure to the sensory, semi-active and active versions of the MAFDS, the mastering of uncertainty is verified and validated within the MAFDS.

3.6.2 *Active Air Spring*

Manuel Rexer and Peter F. Pelz

Suspension systems in automobiles determine the driving comfort for the passengers as well as the driving safety of the vehicle. Currently, two trends are crucial for future suspension systems. Firstly, the trend towards autonomous driving is increasing the demands on driving comfort, as passengers are able to engage in other activities and thus may suffer more frequently from kinetosis [26, 48]. Secondly, the limited electrochemical energy storage of future vehicles requires energy-efficient passive or active suspension systems. This section discusses how both functional requirements can be met by the Active Air Spring, which has been developed and validated over the last 12 years.

It is helpful to understand the function of a suspension system before starting modelling or even designing. The suspension system of a vehicle performs four functions, i.e. (i) carrying the vehicle load, (ii) levelling the distance between the vehicle body and the road, (iii) isolating the body from road or vehicle dynamic excitations and (iv) limiting the dynamic force amplitude of the wheel. These four

functions should be achieved with the least amount of packaging space, weight, energy and cost. Furthermore the four functions should be achieved robustly, i.e. even for uncertain loading or excitation or more general, even for uncertain customer expectations. There are two important quality measures: first, the functional quality as one measure of acceptability; second, the effort measured in energy consumption needed to achieve the functional requirement. Both are addressed in the following section.

Prevailing passive and semi-active systems

For economic reasons, the levelling function (ii) is usually not fulfilled in conventional passive suspension systems. A coil spring enables functions (i) and (iii) with minimal costs. Due to the periodic shift of potential energy from the coil spring with stiffness k_b to the kinetic energy of the chassis of mass m_b and vice versa, the system exhibits a natural frequency $\omega_b = \sqrt{k_b/m_b}$. Above an excitation frequency $\Omega > \sqrt{2}\omega_b$, the road excitations are isolated as desired. To limit resonant oscillations at $\Omega \approx \omega_b$ and to fulfil the functional requirement (iv), namely to limit wheel load oscillation, a hydraulic damper is connected in parallel as a dissipative element. It is immediately apparent that the outlined and prevailing solution is not robust with respect to the uncertainty of the chassis mass m_b . This is due to the non-separate functions (i) and (iii).

Air suspension and hydro-pneumatics are more complex and costly suspension systems. Both enable the levelling function (ii). However, as discussed in Sect. 3.5, true separation of functions as one of seven inherently Robust Design principles is only realised in air suspension. Even with air suspension, the two functions of (iii) isolating the structure from road or vehicle dynamic excitations and (iv) limiting the dynamic force amplitude of the wheel are in conflict, as we show below using a dynamic vehicle model. Only the Active Air Spring, as one example of a smart module as discussed in Sect. 3.5 makes it possible to satisfy new demands resulting from the trends mentioned at the beginning of this section.

Pareto optimal passive and semi-active systems versus Pareto optimal active suspension system

In the case of conflicting tasks, Pareto-optimal solutions emerge. Indeed, we show that the Pareto line cannot be crossed by any active or semi-active state-of-the-art air or hydro-pneumatic suspension systems. In such systems, the spring force $k_b(z, t)z$ and damper force $b_b(\dot{z}, t)\dot{z}$ can be controlled by pneumatic or hydraulic valves. Here, $z = z_w - z_b$ denotes the compression of the suspension system and t the time, cf. Fig. 3.26.

Our research is carried out on three quarter car models, a virtual one, a hardware-in-the-loop system, cf. Sect. 4.3.4, and a real system. The latter is comparable to the MAFDS, Sect. 3.6.1, being a two-mass oscillator as well. The equation of motion of the system depicted in Fig. 3.26 is given by the following two equations of motions and the constitutive equation for the passive force change ΔF_p :

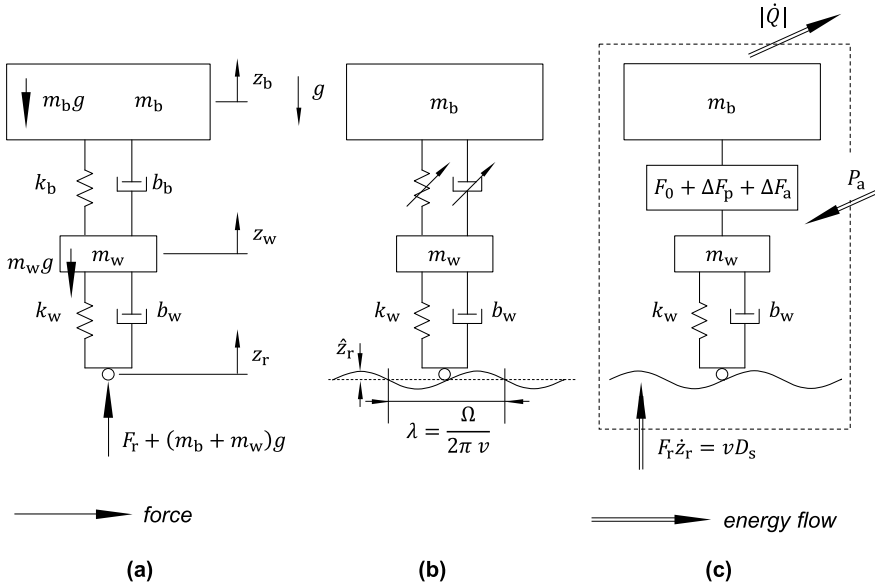


Fig. 3.26 Quarter car model for vertical dynamics of **a** passive, **b** semi-active and **c** active suspension systems with the energy flow across the system boundary

$$m_b (\ddot{z}_b + g) = F_0 + \Delta F_p + \Delta F_a, \quad (3.14)$$

$$m_w \ddot{z}_w + k_w (z_w - z_r) + b_w (\dot{z}_w - \dot{z}_r) = -(\Delta F_p + \Delta F_a), \quad (3.15)$$

$$\Delta F_p = k_b (z_w - z_b) + b_b (\dot{z}_w - \dot{z}_b). \quad (3.16)$$

Here, ΔF_a is the active force between the body and wheel and $F_0 = m_b g$ is the static preload of the suspension system with the mass specific gravity force g . We assume linear springs (stiffness k) and dampers (damping constant b) and a foot point excitation z_r .

The passive, Fig. 3.26a, and semi-active system, Fig. 3.26b, consist of a spring and a damper connected in parallel between body mass, m_b and wheel mass m_w without any active force $\Delta F_a = 0$.

For a semi-active air spring being state-of-the-art today, the stiffness k_b is adapted by increasing the air spring volume V_0 by means of switching pneumatic valves. For semi-active damping systems the damping constant b_b is adapted by switching hydraulic valves.

Here, the focus is on the active force ΔF_a on the right side of the equations. Today there are no Active Air Springs in usage. Figure 3.26c shows the active suspension system. An active system can react to uncertainty from the usage phase and, for example, compensate the loss of comfort as we investigate in Sect. 5.4.5 using the example of the Active Air Spring presented here. When tuning a spring-damper system, the objectives of driving comfort and safety are in conflict because it is not

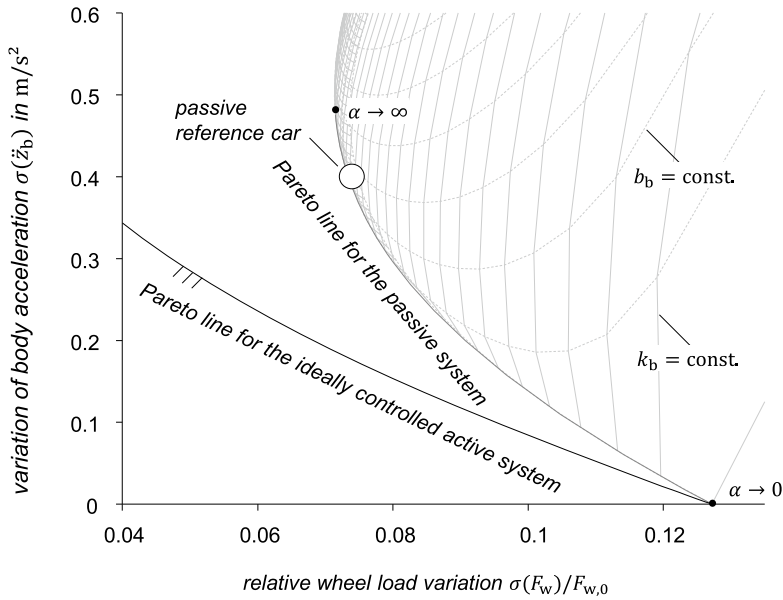


Fig. 3.27 Conflict diagram of driving safety and driving comfort with passive and active boundary lines as well as isolines for constant stiffness and constant damping of a passive suspension for a quarter car driving on a highway at 100 km/h [20]

possible to keep both constant at the same time, the force between the wheel and the ground and the body at rest. Hard-tuned sporty vehicles, for example, offer low driving comfort. As a result, a compromise must be found between the two target parameters during the tuning process. Mathematically, this conflict is described as a minimisation of body acceleration and wheel load fluctuation

$$\min (\sigma(\ddot{z}_b)^2 + \alpha^2 \sigma(F_w)^2). \tag{3.17}$$

The two objectives driving comfort and driving safety—described by the standard deviation of the body acceleration $\sigma(\ddot{z}_b)$ and the standard deviation of the wheel load fluctuation $\sigma(F_w)$ – are weighted via the parameter α .

Figure 3.27 shows the conflict diagram for the quarter car where the standard deviation of body acceleration $\sigma(\ddot{z}_b)$ is plotted versus wheel load fluctuation $\sigma(F_w/F_{w,0})$ with the dynamic wheel load $F_w = k_w(z_w - z_r) + b_w(\dot{z}_w - \dot{z}_r)$ and the static wheel load $F_{w,0} = (m_b + m_w)g$ [34]. Table 3.2 lists the parameter of the passive reference car with a passive linear spring-damper system. The car is excited by a stochastic road signal z_r according to a ride on a highway at 100 km/h [34].

The Pareto line for the passive system, shown in Fig. 3.27 represents the optimal tuning of the spring-damper system of the passive suspension according to Eq. (3.17) with the curve parameter $0 < \alpha < \infty$. This Pareto line cannot be crossed by varying

Table 3.2 Parameters of the quarter car model of the passive reference car according to a middle-range car [20]

Parameter	Value
Body mass m_b	290 kg
Wheel mass m_w	40 kg
Body stiffness k_b	10 000 N/m
Body damping b_b	1140 Ns/m
Tyre stiffness k_w	200 000 N/m
Tyre damping b_w	566 Ns/m

the body stiffness and damping of the system as shown by isolines for body stiffness and damping.

Figure 3.27 also shows the Pareto line for the ideally controlled active system. This system has no limit for the actuating power P_a , cf. Fig. 3.26c. As shown, the functional quality can be significantly improved by an active system. Nevertheless, there is still a conflict between the two targets driving comfort on the one hand and driving safety on the other hand when tuning this system. It is not possible to keep body and wheel at rest at the same time while the system is externally excited. But still, the use of an active system makes it possible to improve both driving comfort and safety and thus to overcome the limits set by the Pareto line of the passive system. In addition, active systems have an extended working range compared to passive or semi-active systems. Due to their variability in force setting they are well suited in mastering the mentioned uncertainty.

Sustainable active suspension demands low viscous or Coulomb friction

From Newton’s third law “actioni contrariam semper et aequalem esse reactionem” [35], known in short as ‘actio est reactio’, it follows that the axial compression force of the air spring is given by

$$F(t) = [p(t) - p_e]A(t). \tag{3.18}$$

This equation applies, provided that the pressure distribution within the component is homogeneous, which is usually the case [38]. In any case, Eq. (3.18) applies if the static pressure $p(t)$ is the volume-averaged pressure in the component and p_e is the ambient pressure. For a piston or plunger, the load-bearing area $A(t)$ is equal to the plunger cross-sectional area. For a rolling lobe or bellow, the load-bearing area $A(t)$ is bounded by a closed curve along which the stress vector has no component in the compression direction, cf. Fig. 3.28. For the Active Air Spring in focus here, the load-bearing area does depend on time t . It is obvious that this can only be achieved by a compliant component. The innovative core of the Active Air Spring presented here is the active inner support of the bellow, cf. Fig. 3.28b.

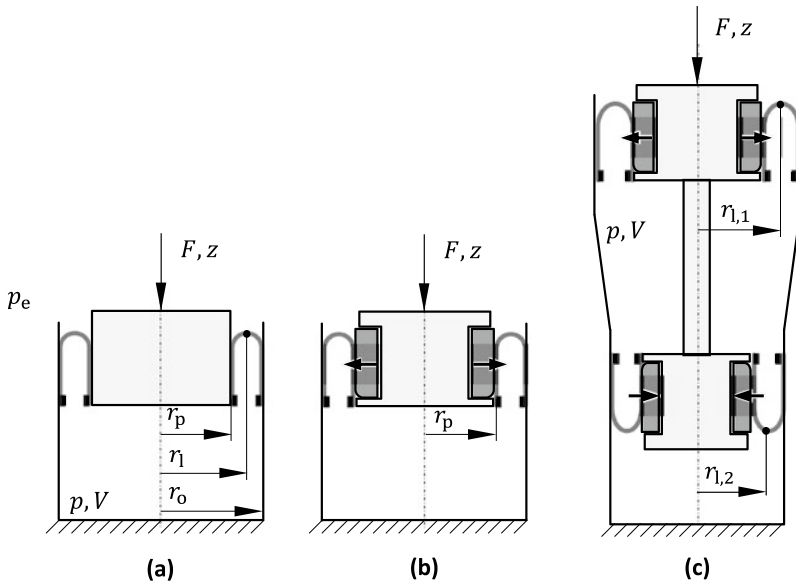


Fig. 3.28 Schematic sketch for **a** bellow sealed air spring with radius of the load-bearing area r_l , **b** actuated piston and **c** double bellow Active Air Spring [20]

For the system at rest at the design point, the spring force is in equilibrium with the gravitational force of the body mass m_{bg} : $F_0 = (p_0 - p_e)A = m_{bg}g$. The index 0 marks the design point.

Transiently, the pressure $p(t)$ within the component can change due to a compliance C , a resistance R or an inductance L or a combination of those effects. In a pneumatic or hydro-pneumatic suspension, the above-mentioned load-bearing function including the levelling function is fulfilled by the fluidic component. The pressure p is adjusted quasi-stationarily for load levelling, cf. Sect. 3.4. This is made possible by a gas compressor in the case of an air spring, and by a displacement pump in case of hydro-pneumatics.

In the Active Air Spring presented here and in Sect. 5.4.5, the gas compliance C enables the second function mentioned above, the energy storage function. Inherent (passive) damping R makes every active system energetically inefficient since the actuator has not only to supply the required energy to the system, but also has to overcome internal frictional, i.e. damping forces. Thus, in the following, damping forces, viscous or Coulomb, are usually considered parasitic in the context of the Active Air Spring and Fluid Dynamic Vibration Absorber (FDVA) introduced in Sect. 5.4.4.

Passive, semi-active, and active pneumatic suspension system

The total force given by Eq. (3.18) can be split into three parts $F(t) = F_0 + \Delta F_p(t, F_0) + \Delta F_a(t, F_0)$. The force $F_0 = m_b g$ conditions the load-bearing function of the suspension system. The passive force change $\Delta F_p(t, F_0)$ is mainly conservative, i.e. results in transient storage of potential energy. If damping is required, e.g. for a passive or semi-active system, the passive force change can also be partially dissipative. For a coil-spring or hydro-pneumatic suspension system, the passive force change $\Delta F_p(t, F_0)$, i.e. the energy storing function, depends on the preload F_0 of the system. The load-bearing and energy-storing functions are therefore not separate. According to Sect. 3.5, an inherently Robust Design concept is the separation of functions, cf. also 3rd case study in Sect. 3.5. This demanded separation of functions is only achieved in the case of air suspension with load levelling:

$$F(t) = F_0 + \Delta F_p(t) + \Delta F_a(t). \quad (3.19)$$

The difference between the active suspension on the one hand and passive or semi-active suspension on the other hand, as discussed in Sect. 3.3, becomes clear when looking at the energy flux per unit time across the system boundary marked by a broken line in Fig. 3.26c.

Three energy flows must be taken into account. Firstly, the mechanical work $P_a = P_a(\Delta F_a)$, which acts on the suspension system per unit time through the actuator of the active component. The power P_a can be positive or negative, i.e. supplied or extracted. Secondly, the excitation work per unit time $P_r = P_r(F_r) = v D_s$, which results from the movement of the vehicle along a bumpy road (amplitude of the waves \hat{z}_r) with velocity v . In the latter case, the energy is taken from the drive system, which is needed for overcoming the additional drag force D_s . This force is additive to the rolling resistance of the wheels, which is essentially determined by viscoelastic-plastic deformation within the wheel elastomer and between the tyre and the road surface, and by the air resistance. Thirdly and finally, the energy dissipated in the damper per unit time is $P_d(F_d) = -|\dot{Q}|$. This power is always negative due to the second law of thermodynamics. It results in a heat flux \dot{Q} . Thus, the first law of thermodynamics for any load-bearing system reads

$$P_a(\Delta F_a) + P_r(F_r) = |\dot{Q}| \Rightarrow \begin{cases} P_r(F_r) = |\dot{Q}| & \text{passive/} \\ P_a(\Delta F_a) + P_r(F_r) = |\dot{Q}| & \text{semi-active,} \\ & \text{active.} \end{cases} \quad (3.20)$$

Equation (3.20) is the foundation for classifying active systems on the one hand and passive or semi-active systems on the other hand, cf. Sect. 3.4. From the first energy balance $v D_s = P_r(F_r) = |\dot{Q}|$ it is easy to derive an upper limit for the energy consumption due to the (a) passive or (b) semi-active suspension system. Neglecting dissipation due to the tyre, for both cases the energy equation reads

$$vD_s = \int_0^{2\pi} k_w(z_r - z_w)\dot{z}_r d(\Omega t) = \int_0^{2\pi} b_b \dot{z}^2 d(\Omega t). \quad (3.21)$$

The angular excitation frequency $\Omega = (2\pi v)/\lambda$ is determined by the wave length λ of the road waviness and the driving velocity v . To simplify the calculation, we assume $m_w \ll m_b$. Therefore, the relative compression can be approximated by $\dot{z} = \dot{z}_w - \dot{z}_b \approx \dot{z}_w$. For this case, the last integral, i.e. the power loss is given by $|\dot{Q}| = \pi b_b \dot{z}_w^2 \Omega^2$. The first integral, i.e. the work exerted by the road on the suspension system, is in any case less than or equal to $\pi k_w \hat{z}_w \hat{z}_r \Omega$. The power needed for moving a vehicle along a bumpy road is

$$D = D_s + \text{air drag} + \text{rolling resistance}, \quad D_s \leq \frac{\pi k_w^2 \hat{z}_r^2}{v b_b}. \quad (3.22)$$

Both air drag and rolling resistance increase with v , while the resistance due to the suspension system decreases with v . This is because the passive damper becomes dynamically stiffer with increasing speed. In the asymptotic limit of $\Omega \rightarrow \infty$, the suspension system becomes energetically conservative: energy storage takes place in the tyre alone.

If only the energy efficiency is taken into account, the damper would have to be replaced by a rigid bar. This is of course nonsensical, as the conflict diagram, Fig. 3.27, indicates. There, the two previously ignored functions of a suspension system are given as coordinates. These are the relative wheel load variation as a measure of driving safety and the body acceleration as a measure of driving comfort.

Before we arrive at that point, we discuss the difference between (a) a passive and (b) a semi-active system both sketched in the above Fig. 3.26. For both systems there is no active force, $\Delta F_a \equiv 0$. For (a) a passive system, the parameters b_b, k_b, b_w, k_w of the ordinary differential equations, Eqs. (3.14)–(3.16), are constant. For a (b) semi-active system, the parameters $b_b(t), k_b(t)$ may vary with time. A semi-active damper or a semi-active air spring enables this. For the first case, the damping constant is changed by adapting the valve opening. For the second case, the spring constant is adapted by adapting the gas volume V_0 without performing mechanical work on the gas. The adaption is reached by opening or closing a pneumatic valve connecting adjacent volumes. For a cyclic excitation, the stiffness is $k_b(t) = k_b(t + T)$, were $T = 2\pi/\Omega$ is the cycle time. Hence, for the semi-active air spring, the suspension system is described by the Hill and Mathieu differential equation [18]

$$m_b \ddot{z} + k_b(t)z = 0. \quad (3.23)$$

It is important to emphasise that the body movement can only be ideally isolated from the movement of the wheel with (c) an active system. This may be achieved by dynamically balancing the spring force (and by balancing a parasitic damper force) with the controlled active force $\Delta F_a(t)$. From Eqs. (3.14) to (3.16) it follows that if the active force is controlled according to the rule $\Delta F_a(t) = k_b z = k_b(z_w - z_b)$,

there will be an ideal isolation, i.e. $\ddot{z}_b = 0$. Ideal isolation becomes more and more important for future autonomous vehicles.

The additional drag D_s , present in the passive and semi-active system, can be reduced by an active system. In an ideal active system there is no damper. Hence, $|\dot{Q}| \rightarrow 0$ and $D_s \rightarrow 0$ for the ideal active system. In the following we analyse possibilities given by air suspension or hydro-pneumatic suspension.

Active and passive means of changing the compression force

To separate active and passive physical effects of changing the compression force, we take a look at the logarithmic derivative of the compression force F , Eq. (3.18),

$$\frac{dF}{F} = \frac{dp}{p - p_e} + \frac{dA}{A}. \quad (3.24)$$

Replacing the increment d by the difference Δ at the design point, we get

$$\frac{\Delta F}{F_0} \approx \frac{\Delta p}{p_0 - p_e} + \frac{\Delta A}{A_0}. \quad (3.25)$$

In general, for the active system we have in Equation (3.24) $dp = dp_p + dp_a$ and $dA = dA_p + dA_a$. Only the active adjustments requires external energy. The passive adjustments for both pressure and load-bearing area are resulting from the relative compression $z = z_w - z_b$ of the spring. A contoured support of the rolling lobe either in form of contoured piston or a contoured outer guidance, results in $A(z)$, cf. Fig. 3.28. For instance, the load-bearing area is by predefined kinematics a function of the compression z . It follows that

$$dA = dA_p + dA_a = \left. \frac{\partial A}{\partial z} \right|_a dz + dA_a = A' dz + dA_a. \quad (3.26)$$

Analogously, the change in gas pressure dp may have a passive and active part: $dp = dp_p + dp_a$. In an asymptotic and hence simplified model, the change in pressure may be isentropic, $s = \text{const}$, or isothermal, $\vartheta = \text{const}$, cf. Sect. 4.3.5. The former is a good model for ‘fast’ processes, i.e. the cycle time T must be much shorter than the thermal relaxation time of the gas. The latter is a good model for ‘slow’ active systems. The load-levelling function of an air spring enabled by a compressor is such a slow active system. Here, there is sufficient time for thermal relaxation of the gas to the ambient temperature. It should be emphasised that only for preliminary design studies, the isentropic or isothermal asymptotes are needed. For observer models the thermal relaxation can easily be calculated, cf. Sect. 4.3.5.

For both, the ‘fast’ and ‘slow’ process, the change of the thermodynamic state is barotropic, because for both cases the pressure is only a function of the gas density: $p = p(\varrho, s = \text{const})$ or $p = p(\varrho, \vartheta = \text{const})$. The volume averaged pressure p is hence only a function of gas volume V and the gas mass m . The gas mass can only be changed (iii) actively by means of a compressor transporting gas from the ambient

into the pressure chamber: $dm = dm_a$. The gas volume is changed (i) passively when the air spring is compressed; the gas volume may be changed (ii) semi-actively by opening or closing a valve connecting the gas volume to an additional gas volume; finally, the gas volume may be changed (iii) actively by means of moving a piston against the gas pressure within the air spring. The passive and active means sum up to $dV = dV_p + dV_a$. With the displacement area A_d defined as $A_d := -dV/dz$ the pressure change is given by

$$dp = \left. \frac{\partial p}{\partial V} \right|_{m,s} (-A_d dz + dV_a) + \left. \frac{\partial p}{\partial m} \right|_{V,\vartheta} dm_a. \quad (3.27)$$

From this, the logarithmic derivative, Eq. (3.24), of the compression force, Eq. (3.18), becomes

$$\begin{aligned} \frac{dF}{F} = & \left(-\frac{A_d}{p - p_e} \left. \frac{\partial p}{\partial V} \right|_{m,s} + \frac{A'}{A} \right) dz + \frac{1}{p - p_e} \left. \frac{\partial p}{\partial V} \right|_{m,s} dV_a \\ & + \frac{1}{p - p_e} \left. \frac{\partial p}{\partial m} \right|_{V,\vartheta} dm_a + \frac{dA_a}{A}. \end{aligned} \quad (3.28)$$

For the design point we have $A_d \rightarrow A_{d,0}$, $A \rightarrow A_0$, $A' \rightarrow A'_0$, $p \rightarrow p_0$, $F \rightarrow F_0 = p_0 A_0$, $m \rightarrow m_0$. Hence, for the ‘fast’ change of the gas volume and the ‘slow’ change of the gas mass we have the asymptotic limits according to Sect. 4.3.5

$$\left. \frac{\partial p}{\partial V} \right|_{m,s} \rightarrow -\gamma \frac{p_0}{V_0}, \quad \left. \frac{\partial p}{\partial m} \right|_{V,\vartheta} \rightarrow \frac{p_0}{m_0}. \quad (3.29)$$

Here, the isentropic exponent is denoted by γ . Hence, the logarithmic derivative becomes

$$\frac{\Delta F}{F_0} = \frac{\Delta F_p}{F_0} + \frac{\Delta F_a}{F_0} = k_b \frac{\Delta z}{F_0} + \frac{\Delta F_a}{F_0}, \quad (3.30)$$

with the abbreviation for the stiffness

$$k_b := \frac{\Delta F_p}{\Delta z} = m_b g \left(\gamma \frac{p_0}{p_0 - p_e} \frac{A_{d,0}}{V_0} + \frac{A'_0}{A_0} \right) \quad (3.31)$$

and the active force contribution

$$\frac{\Delta F_a}{F_0} = \frac{p_0}{p_0 - p_e} \left(\frac{\Delta m_a}{m_0} - \gamma \frac{\Delta V_a}{V_0} \right) + \frac{\Delta A_a}{A_0}. \quad (3.32)$$

Inherent robust design concept of the Active Air Spring

From the stiffness, Eq. (3.31), we recognise the independence of the body natural frequency from the body mass m_b for $p_0 \gg p_e$ (this condition is usually fulfilled)

provided V_0 is controlled to be constant by means of a levelling function, i.e. a compressor as actuator,

$$\omega_b := \sqrt{\frac{k_b}{m_b}} = \sqrt{g \left(\gamma \frac{A_{d,0}}{V_0} + \frac{A'}{A_0} \right)}. \quad (3.33)$$

Thus, the system is inherently robust due to the separation of the two functions, first storage of potential energy and second levelling the body relative to the wheel to a predefined distance. As such, it is self-adaptive to load changes, meaning it compensates for uncertainties in the load. This is the reason why air suspension is standard in commercial vehicles and rail cars. The three principles of ‘separation of functions’, ‘self-adaptation’, and ‘compensation’ are three of the seven inherently Robust Design concepts discussed in Sect. 3.5.

From Eqs. (3.20) to (3.25) we recognise the semi-active nature of air springs with several volumes: by adapting the volume V_0 , the stiffness and by that the dynamic behaviour is changed. But the semi-active control fails in preventing motion sickness. Thus, the following section discusses how to achieve an active force ΔF_a .

For the active system, the relevant cycle time is $T = 2\pi/\omega_b \sim 1$ s. Changing the gas mass Δm_a in a comparable time is not feasible. Hence, changing the gas mass is limited to the levelling function of the suspension system. Dynamically adapting the gas volume ΔV_a may be achieved by pumping oil, e.g. by a gear pump, as is done for hydro-pneumatic suspension systems. For an air suspension system to become an Active Air Spring, only the adaption of the load-bearing area is feasible [6]. Hence, for ‘fast’ cyclic forces $\Delta F_a(t) = \Delta F_a(t + T)$. Equation (3.32) reduces to

$$\Delta F_a = p_0 \Delta A_a. \quad (3.34)$$

Milestones in 12 years development of the Active Air Spring

Figure 3.29 shows the milestones in twelve years of development of the Active Air Spring as a time line. From the idea, the development led via a first prototype, which realises the adjustment of the load-bearing area via a hydraulically driven cam gear [4] to the second prototype with two connected hydraulic membrane actuators [20]. We were able to significantly reduce both the package and the weight of the prototype. The result is a hydraulic diaphragm actuator that is completely integrated into the piston of the Active Air Spring (Fig. 3.32). To ensure the relative force change $\Delta F_a/F_0$ is as large as possible, a double bellows air spring is used. Figure 3.28c shows the structure of two pistons connected by a rod. The load-bearing area thereby describes a circular ring.

Both pistons are equipped with an actuator allowing a total force change of $\Delta F_a \approx \pm 1$ kN at a static load of $F_0 = 2850$ N. A further advantage of the double bellow Active Air Spring is the hydraulic coupling of the two actuators via a double acting cylinder. As a result, the energy requirement of the air spring is reduced, as only differential forces have to be set, comparable to a mechanical rocker [43]. Fur-

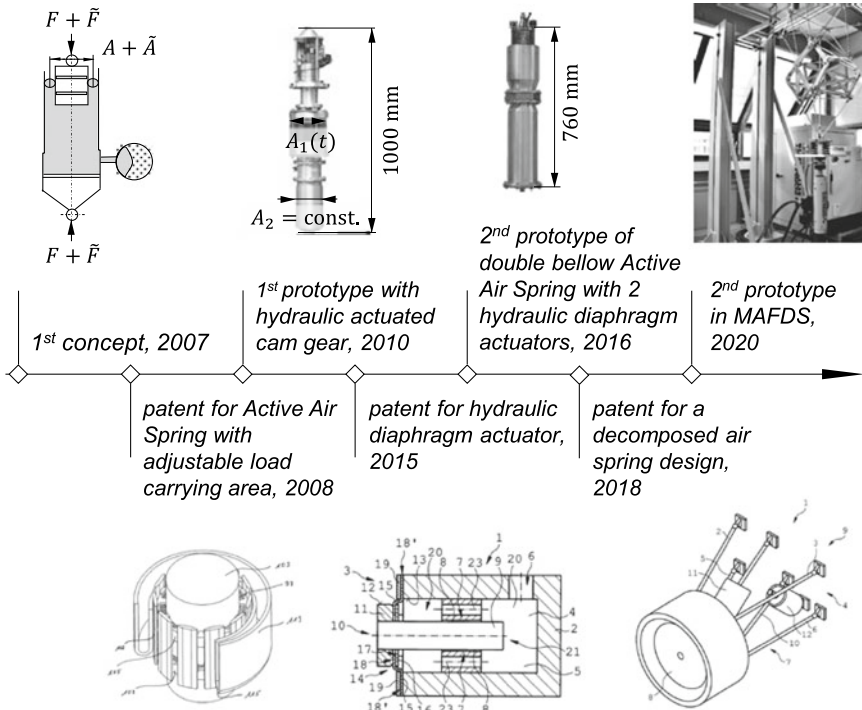


Fig. 3.29 Milestones in the development of the Active Air Spring [21, 22, 39]

thermore, energy can be recuperated by resetting the actuators through the bellows. Figure 3.30 shows the concept of the Active Air Spring including two actuator pistons and the hydraulic coupling. The actuator can be driven by different operating principles. A hydraulic drive is used for the technical prototype.

Table 3.3 gives the main characteristics of the developed Active Air Spring spring. Figure 3.31 shows the designed second prototype according to these parameters.

The two actuators are the core of the Active Air Spring. They work as single-acting hydraulic linear actuators. The reset is done via the bellows and the pressure in the air spring. In order to distribute the load on the bellows, the piston is divided into four segments [5], each moves radially up to ± 3 mm [22]. These segments are guided by piston rods which are mounted in the piston via sliding bushes. Diaphragm cloths seal the actuators with low friction and without leakage [22]. Figure 3.32 shows the structure of the upper actuator.

Influencing factors on the performance of the active suspension system are

1. the maximum velocity at which the actuator forces can be adjusted,
2. the maximum actuator force $\Delta F_{a,max}$ and
3. the control concept and the selected controller.

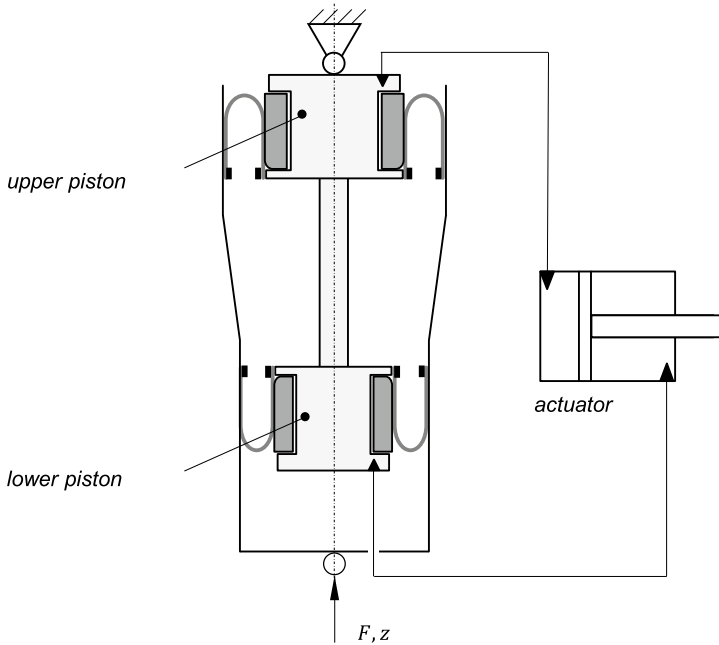


Fig. 3.30 Double bellow Active Air Spring with connected hydraulic actuators [20]

We considered all these factors in the development of the technical prototype. The uncertainty in the design parameters, such as the body mass for example, was taken into account in the sense of a resilient product development [25].

In the following, we present the results necessary for an optimal design and operation of the Active Air Spring.

1. *Actuator Velocity.* Active suspension systems can be divided in slow active systems working up to 5 Hz and fast active systems working up to 30 Hz. We showed that a slow active system is sufficient to improve driving comfort and avoid kinetosis [20]. The transfer function of the developed actuator corresponds approximately to a first order low pass filter with a cut-off frequency of 5 Hz [23].
2. *Actuator Force.* We showed that an actuator force of about 1 kN is sufficient to influence the vertical dynamics of a car with the properties shown in Table 3.2 [20]. With more available force, the driving comfort is not further improved.
3. *Controller.* The control concept consists of a primary controller that specifies the actuator force and a secondary controller that controls the realisation of the required actuator force via the hydraulic diaphragm actuators [20]. The structure and parameters of the primary control loop determine the performance of the active system. We designed the controller via an H2-optimisation [32]. The resulting controller structure includes a skyhook controller combined with a preview function. The preview function significantly increases the performance of

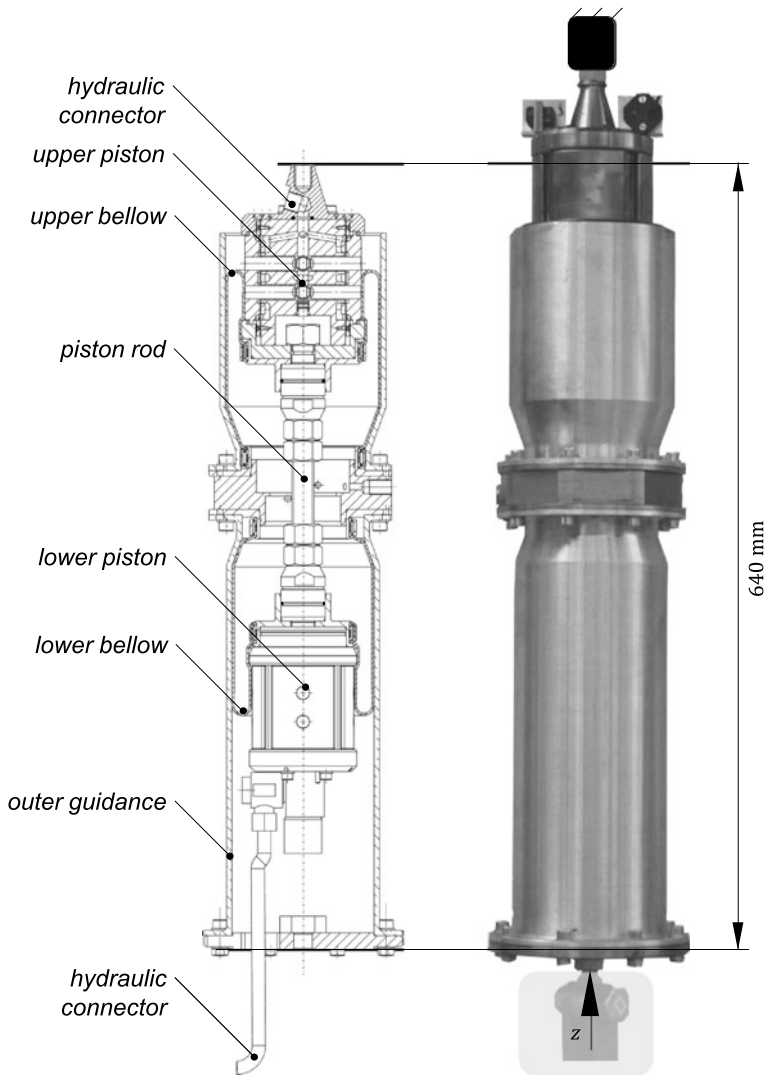


Fig. 3.31 Technical prototype of the Active Air Spring [20]

the active system. This results in almost the ideal active system [32]. Therefore the road excitation z_r with velocity \dot{z}_r and acceleration \ddot{z}_r as well as the body velocity \dot{z}_b is fed back statically. More complex dynamic controllers are not used due to stability reasons [24]. Table 3.4 lists the implemented controller parameters. The designed controller is robust against uncertain parameters such as varying body mass m_b or wheel stiffness k_w . Figure 3.33 shows that the results are always close to the active Pareto line. The controller is designed with the nominal con-

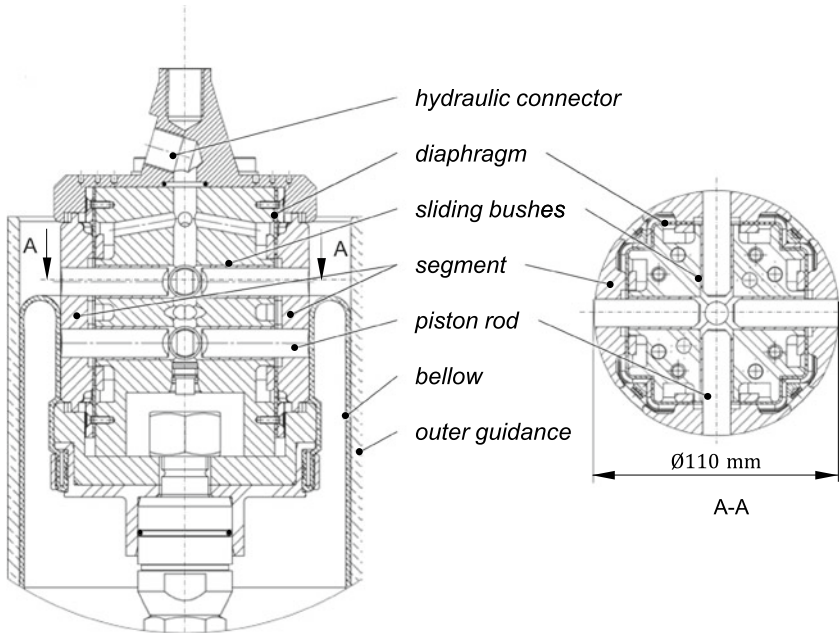


Fig. 3.32 Sectional view of the hydraulic diaphragm actuator with four segments [20]

figuration shown in black. The Pareto lines for active systems are recalculated for each configuration as a reference. The controller is also robust against uncertainty in excitation as shown in Sect. 5.4.5 and [24].

The example of the Active Air Spring is used in several chapters throughout this book. The possible improvements in driving comfort with the active suspension system are presented in Sect. 5.4.5. These experimental investigations were realised on a hardware-in-the-loop test rig. The model uncertainty of these tests is shown in Sect. 4.3.4. The actuator is also used as an example of a resilient process chain (see Sect. 6.3.7).

3.6.3 3D Servo Press

Maximilian Knoll, Florian Hoppe, and Peter Groche

Forming machines are used to provide forming forces and energies required to form and guide the tools. Depending on the requirements, individual types of machines are used. Due to fluctuations on the downstream market, the requirements for products and thus the forming processes can vary tremendously. As forming machines come with a high investment, it is important to predict future requirements and select

Fig. 3.33 Simulated influence of uncertain system parameters on the Active Air Spring when using the skyhook controller with preview (square) when riding on a highway with 100 km/h [20]

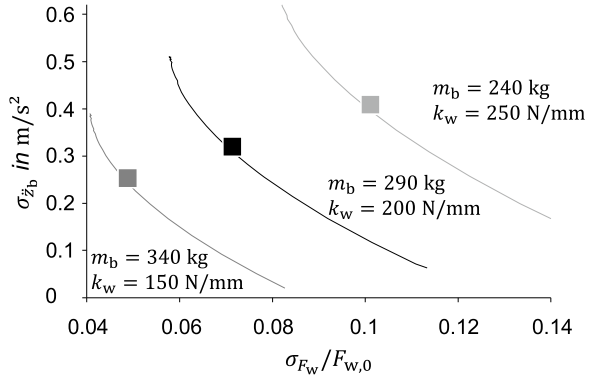


Table 3.3 Important characteristics of the Active Air Spring [20]

Parameter	Value
Static load F_0	2850 N
Actuator force $\Delta F_a(z = 0)$	± 1180 N
Static pressure p_0	14 bar(a.)
Total load-bearing area A_0	(2195 ± 910) mm ²
e maximum deflection z_{max}	± 70 mm
Air spring volume V_0	2.21
<i>Upper piston</i>	
Piston diameter $d_{p,1}$	(105 ± 6) mm
Outer diameter $d_{o,1}$	140 mm
Load-bearing area A_1	(11.740 ± 570) mm ²
Segment high $h_{seg,1}$	76 mm
<i>Lower piston</i>	
Piston diameter $r_{p,2}$	(94 ∓ 5) mm
Outer diameter $r_{o,2}$	127 mm
Load-bearing area A_2	(9545 ∓ 340) mm ²
Segment high $h_{seg,2}$	76 mm

Table 3.4 Implemented skyhook controller with preview

Parameter	Value
k_{z_b}	-1750 Ns/m
k_{z_r}	-3860 N/m
$k_{\dot{z}_r}$	-1016 Ns/m
$k_{\ddot{z}_r}$	-18.29 Ns ² /m

an optimal machine type. Usually the machine types are divided into force-driven, path-driven and energy-driven machines. All of them have in common that they move a press ram which guides the tool along a horizontal or vertical axis. While energy-driven machines, e.g. hammers, provide a defined forming energy in terms of kinetic ram energy, force-driven machines, e.g. hydraulic presses, provide a specific maximum force over the complete ram motion. The actual ram motion is then defined by the input energy or force and the reacting forces. In contrast, a conventional path-driven machine consists of an electric drive, a clutch, flywheel and crank drive; therefore it provides a predetermined pattern of ram motion [54]. The non-modifiable pattern of ram motion (e.g. sinusoidal) limits the fields of application.

By integrating servo drives, new types of presses have been introduced to overcome these limitations. These servo presses allow the ram motion to be adjusted by means of control systems and algorithms. However, these machines so far only allow ram movements in one direction. To achieve more demanding geometries to make use of properties difficult to deform, special machines or complex tool designs are often unavoidable.

One prominent case of special processes is orbital forming. During orbital forming a tool with small contact areas rolls over the workpiece and progressively shapes it [8]. Forming forces are reduced drastically in orbital forming presses, the translatory ram motion is extended by an additional rotary motion. While the translatory ram motion can be freely programmed, the rotary motion can be adapted to the process only in restricted levels.

Classical path-driven presses with a one-dimensional motion are designed for high productivity and have proven their value over the last decades. Especially path-driven presses with flywheel are able to produce a huge number of parts per minute. But they are limited to one degree of freedom ram motion. Today's servo presses provide an increased flexibility, as described in Sect. 6.21.

But they are not able to offer motion patterns which special presses use to get the most out of the individual forming process.

As seen from the previous paragraphs, classical and special presses either offer a high productivity or specialised motions. After an investment decision for a forming machine is made, the flexibility with respect to motion pattern and productivity is very limited. But due to uncertain conditions and requirements on the upstream and downstream market, modern production systems should be able to switch over to a multitude of processes and process chains. This is only possible through the cost-intensive acquisition of different machine types.

Furthermore, disturbances that occur during production, such as varying material properties, are difficult to control. Since existing machine concepts only offer very limited possibilities to compensate for these, tight tolerance requirements are set for material and process parameters which are continuously monitored. Yet, increasing the adaptability of the process has been accompanied by productivity reductions and the installation of additional drive systems.

Our objective is therefore to combine the advantages of different press types and to enable adaptability without loss of productivity. A new approach is the independent actuation of the press ram at three points [16]. An independent motion of these

points allows any desired motion of the ram in three degrees of freedom (dof). While mechanical presses highly rely on a stiff design to achieve a high accuracy, a free 3-dof ram motion allows for a compensation of mechanical inaccuracies, elasticities and disturbances by means of control systems [17]. Thus it is possible to break away from previous design constraints and to pursue new design objectives.

Scheitza [45] proved the feasibility of these theoretic considerations of a new 3D Servo Press concept with a mechanical demonstrator with 1 ton press force, Fig. 3.34. Each of the three ram points is driven by an independent press gear which consists of eccentric and modified knuckle joint kinematics. To adapt the shut height of the tools, two additional spindle kinematics modify the press gear. Several studies have shown that the new press concept allows to extend process limits and combine processes, thus increasing flexibility and productivity. The ability to react to disturbances paved the way to control product properties, see Sect. 5.3.2.

After successful validation of the 3D Servo Press concept, it was scaled up from a prototype with 1 ton press force to a 160 ton press, as shown in Fig. 3.34. While the press force is upscaled by a factor of 160, geometrical dimensions are only upscaled by a factor of approximately 4. This leads to a number of challenging size effects. The rolling bearings used in the prototype reach a technical limit in terms of load capacity and dimension at this scale. Plain bearings normally used in press design are accompanied by bearing clearance, which adds up through the kinematic chain. Investigations on this scaling effect led to the development of novel combined roller and plain bearings, as described in Sect. 5.4.2. With the integration of such a combined bearing design the bearing size has been reduced, while at the same time increasing the service life and stiffness of the bearing [47].

Scaling the ability to mount larger tools comes with a dimensional increase in the ram size. But to achieve the same tilting angle with the 1600 kN press, the ram drive points have to cover a larger distance, which is called the stroke height. Therefore, the stroke height had to be scaled up from 41 to 100 mm with an additional adaption of the shut height of 200 mm. Besides technical constraints and scalability laws, also requirements and restrictions by the infrastructure department had to be complied with when scaling up the press. Due to the associated maximum installation area and height, linear scaling of the press kinematics proved to be not practicable. In order to realise the three gears in the required dimensions, a new arrangement of individual levers was carried out, compare Fig. 3.34. The scaling of the stroke height results in a smaller width to height ratio.

During the scaling of the gear dimensions the absolute manufacturing tolerances increase, adding up in the kinematic chain. These have a significant effect on the accuracy of the ram positioning and thereby on product quality [17]. Challenges are to produce individual parts of up to $2000 \text{ mm} \pm 20 \mu\text{m}$ in length with homogeneous material properties.

Next to manufacturing individual parts, assembly and delivery bring up new scaling challenges. Due to infrastructural limitations, the press could not be fully assembled at the operating site. Developments of innovative products come with a significant amount of uncertainty, especially with respect to functionality. This requires strategies to reduce both uncertainty and cost. Therefore, the press had to be assem-

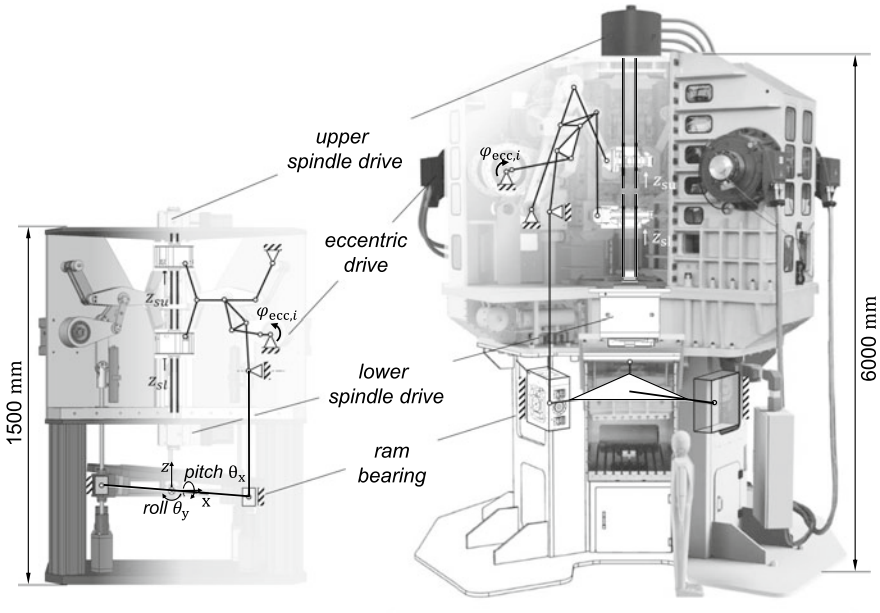


Fig. 3.34 Kinematic comparison of 1 and 160ton 3D Servo Press

bled at an assembly location where also all functions had to be tested. To reach a cost minimum between transportation costs and assembly hours, the press was designed to be disassembled into two parts, transported to the operating site and reassembled with special equipment within a few hours.

The electromechanical devices had to be scaled to drive the scaled gear mechanisms and provide the pressing force. Presses require high torques but typical electrical machines provide high velocities at low torque. The prototype was equipped with gearboxes to increase the torque transmission, but that comes with the disadvantages of reduced efficiency, lower stiffness, large installation place, increased weight and maintenance. Hence, a gearbox cannot be upscaled for the 160 ton press; thus direct drive design is desired which requires high-torque drives. The diameter of high-torque drives is significantly larger than those of typical drives, as the torque in electrical machines scales linearly in stator length and quadratically in stator diameter. To achieve a symmetric torque on the eccentric shaft, the 160 ton press was equipped with 2 torque drives on each eccentric axis, the shafts of which being positively coupled with the eccentric.

Typical ram motions such as a press stroke or an orbital motion can be adopted from state of the art press control in which the drive motion defines the ram motion. But the 3D Servo Press is designed to control the actual product properties as described in Sect. 5.3.2. Therefore it allows for arbitrary 3D motions of the ram, which requires a paradigm shift in the control. Control methods can be adopted from robot control in which the drives are controlled in a closed-loop process and the kine-

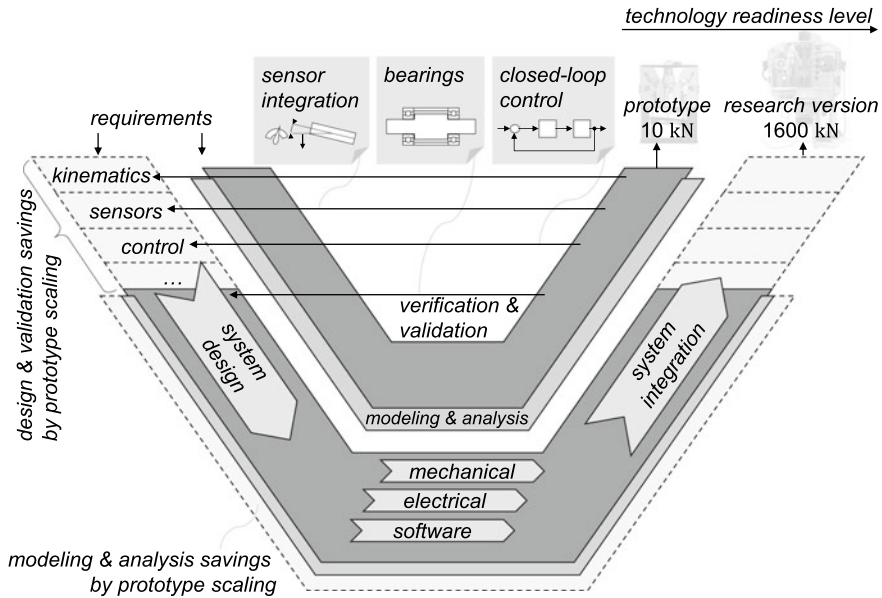


Fig. 3.35 Scaling of development method and design of the 3D Servo Press using the V-model

matic is controlled in an open-loop by means of inverse kinematics [28]. First control approaches were investigated in simulations, being implemented and validated on the prototype. This method corresponds to the V-model shown in Fig. 3.35, based on feeding back experience in the validation phase to the design phase. As lack of knowledge plays a major role in innovative research projects, early validation phases are crucial to be able to adapt to new insights. By performing time-consuming design and validation steps on the prototype, these steps can be significantly reduced for the 160t 3D Servo Press whereby commissioning time and risk can be cut.

Early validation phases have shown that since the limited installation space did not allow a press design with maximum mechanical stiffness, the press elasticity has a greater effect on accuracy. Hence, the lower passive stiffness has to be increased by means of stiffness models and active measures, i.e. a closed-loop control of the kinematics as described in Sect. 5.4.1. A closed-loop control of the kinematics requires real-time adaption of the drive motion and hence real-time solution of inverse kinematics. As no explicit solution is available for the inverse spindle and ram kinematics, real-time optimisation routines were applied but are computationally expensive. But as higher control speed only comes with a higher sampling time [28], the optimisation steps had to be parallelised with the control. Stability analyses were carried out in experiments on the prototype and have demonstrated the effect of uncertainty that inhibits industrial closed-loop ram control today. The early analyses on the prototype also allowed to develop new robust control methods shown in Sect. 6.1.7.

An important aspect of the motion control of a press is the measurement of the actual ram position in three degrees of freedom. While a tactile measurement of the ram would be most accurate, it has to be placed in the working area and therefore is less robust regarding potential damage, see Sect. 5.2.4. On the other hand, visual measurement methods must trade off between sampling rate and accuracy and are affected by oil mist, lighting conditions and others. To receive a robust but also accurate position [7], force and position sensors have been placed in the gear mechanism close to the ram to be combined with an observer model.

An additional force measurement of ram forces pursues two tasks: (A) enable accurate force control and (B) protect the machine against overload. As direct force measurement in the force flow reaches technical limitations when scaling up, only indirect force measurement is applicable. This requires a structure with linear elastic behaviour and continuous cross-section. To reduce unwanted effects, the measurement should be close to the ram. The drive rods that connect gear and ram bearings are highly suitable for that task due to their ideal linearity. But due to its length and geometrical uncertainty, minor buckling is to be expected. Buckling distorts force-measurement if only measured on one face. Therefore three piezo force-sensors were installed on the drive bars. One way to compensate the buckling effect is to calculate the mean of the sensors via software. Another way is to use the piezo sensors as an electric circuit to calculate the mean. By connecting the piezo crystals in series their charge physically sums up; this performs the same function as the mean, however by even tripling the accuracy.

Besides the control system, the majority of the software modules were likewise developed and validated on the basis of the prototype. This includes functions, such as the user interface, sensor evaluation, logic, motion control and real-time communication. While a user-interface is run by an industrial PC communicating with a programmable logic controller (PLC) via non-real-time ethernet, a real-time communication between PLC, sensors and drives is mandatory. This involves also the synchronisation of the positively coupled torque drives. In contrast to the prototype, each shaft that drives an eccentric gear is equipped with two 3500 Nm torque drives. To add up the torques of both drives and prevent that both drives operate against each other, both drives synchronise their torque in a master-slave setup. While the master drive is controlled by the kinematic press controller, the second drive is directly linked to the measured torque of the master and thus supports its motion.

As research tasks changing over time require modifiable software, safety functions must be outsourced to non-modifiable hardware. Hence, safety functions are performed by sensors and an additional safety PLC. Although the design of the functional safety of the 3D Servo Press for compliance with the Machinery Directive 2006/42/EC [12] is based on typical presses, its machine safety has required several innovations.

While the actual press force on one hand serves as a process and control variable, the force sensors also protect against overload. When the maximum force is reached in case of a fault, the machine motion is decelerated to zero. As three degrees of freedom are involved in the motion, a simply upward motion of the ram might damage the tool. The new closed-loop force control allows to safely move the ram up to a

limited force and stop once the tool is locked. In contrast to typical servo presses, the spindle kinematics of the 3D Servo Press is designed to continuously adapt the dead centres during the process. Due to the high kinetic energy in the spindles during the process, multiple adaptations of the initial prototype design have been performed. This involves a large amount of kinetic energy being converted into heat. Therefore, highly efficient spindles have been installed in the 3D Servo Press. However, the converted heat amount depends on the load and motion history, which is why both spindle nuts are additionally monitored with temperature sensors. The second spindle related safety aspect is maintaining the mechanical working area. On the one hand, the mechanical limitation of each individual spindle must be protected. In contrast to the prototype, the spindle force of the upscaled version is larger than the mechanical load capacity and therefore requires a sensor-based end stop. However, due to the high potential kinetic energy, it is impossible to ensure that the spindles decelerate in time when the end stop is reached. Therefore, a second sensor for speed reduction was integrated before the end stop. On the other hand, operating the window nonlinearly depends on both spindle positions. While the control software guarantees to keep these boundary conditions, an additional non-tactile sensor monitors compliance with the process window. To improve the control performance and accuracy, combined roller-plain bearings were integrated in each joint. Those are lubricated via multiple circuits which are being monitored in terms of flow, pressure and temperature.

Due to its 8 servo drives, the 3D Servo Press requires a maximum electric power of 1.2 MW resulting in a nominal mechanical press force of 160 tons at a maximum speed of 100 strokes/min. While eccentric drives contribute to the maximum speed with 100 kW each, the spindle drives come with 300 kW. The spindle kinematics is designed for a high force transmission which allows to freely adapt the gear during the process without being disturbed by reacting forces. But the spindles are only allowed to move in a defined process window which yields in the achievable stroke and shut height. The process window of the spindle positions (Fig. 3.36a) can also be mapped to the shut and stroke height, see Fig. 3.36b). This results in a shut height range from 500 to 700 mm and a stroke height adjustment from 25 to 100 mm, see Fig. 3.36b.

The eccentric kinematics results in a position-dependent force transmission of the eccentric drives and therefore position-dependent maximum ram force. Under the assumption of an inelastic system, an infinite transmission ratio can be reached at the top and bottom dead centre. The nominal force of the 3D Servo Press is reached $\varphi_{ecc,i} = \pm 16^\circ$ before the bottom dead centre. The maximum force reaches its minimum between the two dead centres.

Due to the 120° arrangement of the three gears, which are coupled via the central spindles, the eccentrics can be controlled independently, as shown in Fig. 3.34. As a result, both a translatory stroke and an orbital motion, as well as anything in between can be realised.

The maximum stroke height of 100 mm and the distance between the ram bearings result in a maximum pitch angle of 3.44° and maximum roll angle of 2.97° .

Starting from an initial design of the 1 ton press in 2008, the 160 ton press was developed while starting the production of first parts in 2011, see Fig. 3.37. In the

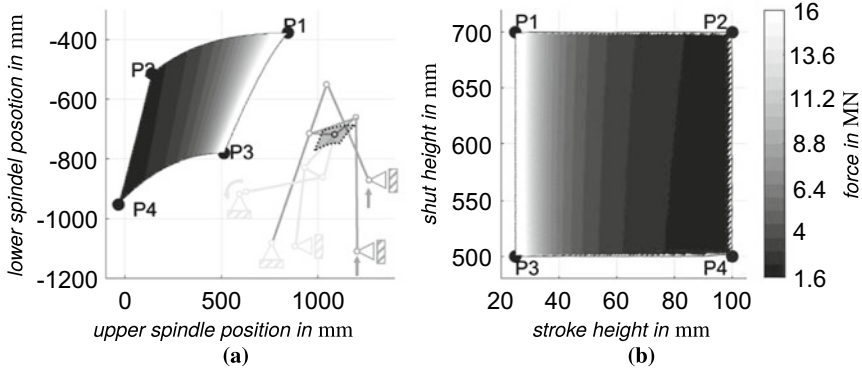


Fig. 3.36 Process window of the 3D Servo Press in combination of minimum process force at maximum engine power, **a** as a function of the spindle positions and **b** transformation to the shut and stroke height

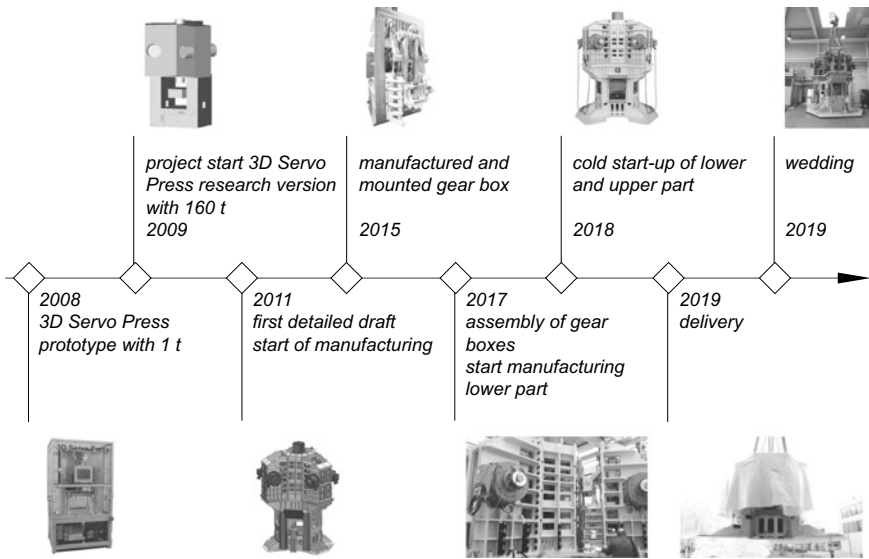


Fig. 3.37 Timeline of development 3D Servo Press from 2008 to 2019

following years, the press was further detailed as the production of individual parts progressed. In 2015, the first major milestone was reached with the assembly of the first gearbox. In the same year, the centrepiece of the 3D Servo Press was realised. The completion of the upper part began in 2017 and was completed with the assembly of the last gear box to form a fully integrated gearbox. Subsequently, the assembly of the lower part and cold commissioning of the press was carried out at the assembly site and assured the functionality of the 3D Servo Press. Following the cold commis-

sioning the press was transported to the operating site, where the final positioning and commissioning took place.

References

1. Ackermann J (2002) Robust control: the parameter space approach, 2nd edn. Communications and control engineering. Springer, London
2. Ashby MF, Jones DRH (2005) Engineering materials. Elsevier/Butterworth-Heinemann, Amsterdam
3. Baucells M, Borgonovo E (2013) Invariant probabilistic sensitivity analysis. *Manag Sci* 59(11):2536–2549. <https://doi.org/10.1287/mnsc.2013.1719>
4. Bedarff T (2017) Grundlagen der Entwicklung und Untersuchung einer aktiven Luftfeder für Personenkraftwagen. Forschungsberichte zur Fluidsystemtechnik, vol 10. Shaker, Herzogenrath. Dissertation, TU Darmstadt
5. Bedarff T, Hedrich P, Pelz PF (2014) Design of an active air spring damper. In: Murrenhoff H (ed) 9th international fluid power conference (9th IFK). HP - Fördervereinigung Fluidtechnik, Aachen, pp 356–365
6. Bretz A, Calmano S, Gally T, Götz B, Platz R, Würtemberger J (2015) Darstellung passiver, semi-aktiver und aktiver Maßnahmen im SFB 805-Prozessmodell. Preprint
7. Calmano S, Schmitt S, Groche P (2013) Prevention of over-dimensioning in light-weight structures by control of uncertainties during production. In: International conference on “New developments in forging technology”. MAT INFO Werkstoff-Informationsgesellschaft mbH, Fellbach, pp 313–317
8. Calmano S, Hesse D, Hoppe F, Traidl P, Sinz J, Groche P (2015) Orbital forming of flange parts under uncertainty. In: Pelz PF, Groche P (eds) Uncertainty in mechanical engineering II, vol 807. Applied mechanics and materials. Trans Tech Publications, pp 121–129
9. de Secondat Montesquieu C (1750) The spirit of laws (Nourse J, Vaillant P in the Strand, London). Translated from the French of M. de Secondat, Baron de Montesquieu
10. Efron B (1979) Bootstrap methods: another look at the jackknife. *Ann Stat* 7(1):1–26. <https://doi.org/10.1214/aos/1176344552>
11. Enss GC, Gehb CM, Götz B, Melz T, Ondoua S, Platz R, Schaeffner M (2016) Device for optimal load transmission and load distribution in lightweight structures (Kraftübertragungsvorrichtung). Patent DE 10 2014 106 858 A1
12. European Parliament (2006) Council of the European Union: Directive 2006/42/EC of the European Parliament and of the council of 17 may 2006 on machinery, and amending directive 95/16/EC. Official J Eur Union. <https://eur-lex.europa.eu/eli/dir/2006/42/oj>
13. Feldmann R, Gehb CM, Schaeffner M, Matei A, Lenz J, Kersting S, Weber M (2020) A detailed assessment of model form uncertainty in a load-carrying truss structure. In: Model validation and uncertainty quantification, vol 3. Conference proceedings of the Society for Experimental Mechanics series. Springer
14. Fowlkes WY, Creveling CM (2012) Engineering methods for robust product design: using Taguchi methods in technology and product development. Engineering process improvement series, Addison-Wesley, Reading, Mass
15. Ghanem R, Higdon D, Owhadi H (2017) Handbook of uncertainty quantification. Springer, Cham
16. Groche P, Scheitza M, Kraft M, Schmitt S (2010) Increased total flexibility by 3D servo presses. *CIRP Ann* 59(1):267–270
17. Groche P, Hoppe F, Sinz J (2017) Stiffness of multipoint servo presses: mechanics vs. control. *CIRP Ann* 66(1):373–376
18. Hagedorn P (1982) Non-linear oscillations, vol 10. The Oxford engineering science series. Clarendon Press, Oxford

19. Hanselka H, Platz R (2010) Ansätze und Maßnahmen zur Beherrschung von Unsicherheit in lastragenden Systemen des Maschinenbaus. *Konstruktion* 2010(11–12):55–62
20. Hedrich P (2018) Konzeptvalidierung einer aktiven Luftfederung im Kontext autonomer Fahrzeuge. Dissertation, TU Darmstadt. <https://tuprints.ulb.tu-darmstadt.de/8469/>
21. Hedrich P, Pelz PF (2018) Einzelradaufhängung für ein Kraftfahrzeug. Patent DE102018127301A1, 31 Oct 2018
22. Hedrich P, Johe M, Pelz PF (2015) Aktor mit einem linear verlagerbaren Stellglied. Patent DE 102015120011 A1, 18 Nov 2015
23. Hedrich P, Lenz E, Pelz PF (2017) Modellbildung, Regelung und experimentelle Untersuchung einer aktiven Luftfederung in einer Hardware-in-the-Loop-Simulationsumgebung. VDI-Fachtagung Schwingungen 2017. VDI-Berichte Band 2295. VDI Verlag, Düsseldorf, pp 447–460
24. Hedrich P, Lenz E, Pelz PF (2018) Minimizing of kinetosis during autonomous driving. *ATZ Worldwide* 120(7–8):68–75
25. Hedrich P, Brötz N, Pelz PF (2018) Resilient product development – a new approach for controlling uncertainty. In: Pelz PF, Groche P (eds) *Uncertainty in Mechanical Engineering III*, vol 885. Applied mechanics and materials. Trans Tech Publications, pp 88–101. <https://doi.org/10.4028/www.scientific.net/AMM.885.88>
26. Heinrichs D (2015) Autonomes Fahren und Stadtstruktur. In: Maurer M, Gerdes JC, Lenz B, Winner H (eds) *Autonomes Fahren: Technische, rechtliche und gesellschaftliche Aspekte*. Springer, Berlin, pp 219–239. https://doi.org/10.1007/978-3-662-45854-9_11
27. Holl M, Rausch L, Pelz PF (2017) New methods for new systems - how to find the technoeconomically optimal hydrogen conversion system. *Int J Hydrog Energy* 42(36):22641–22654. <https://doi.org/10.1016/j.ijhydene.2017.07.061>
28. Hoppe F, Pihan C, Groche P (2019) Closed-loop control of eccentric presses based on inverse kinematic models. *Proc Manuf* 29:240–247. <https://doi.org/10.1016/j.promfg.2019.02.132>
29. Ihn T (2016) Data analysis – Fehlerfortpflanzung. Lecture Notes, ETH Zürich
30. Iooss B, Saltelli A (2016) Introduction to sensitivity analysis. In: Ghanem R, Higdon D, Owghadi H (eds) *Handbook of uncertainty quantification*. Springer, Cham, pp 1–20. https://doi.org/10.1007/978-3-319-11259-6_31-1
31. Isermann R (2008) *Mechatronische Systeme: Grundlagen*, 2nd edn. Springer, Berlin
32. Lenz E, Hedrich P, Pelz PF (2018) Aktive Luftfederung – Modellierung, Regelung und Hardware-in-the-Loop-Experimente. *Forschung in Ingenieurwesen* 82:171–185. <https://doi.org/10.1007/s10010-018-0272-2>
33. Miner M (1945) Cumulative damage in fatigue. *J Appl Mech* 12:159–164
34. Mitschke M, Wallentowitz H (2014) *Dynamik der Kraftfahrzeuge*, 5th edn. Springer, Wiesbaden
35. Newton I (1726) *Philosophiae naturalis principia mathematica*. Tomus Primus, London
36. Pahl G, Beitz W, Feldhusen J, Grote KH (2007) *Engineering design: a systematic approach*, third edn. Springer, London. <https://doi.org/10.1007/978-1-84628-319-2>
37. Palmgren A (1924) Die Lebensdauer von Kugellagern. *Verein Deutscher Ingenieure* 68:339–341
38. Pelz P, Buttenbender J (2004) The dynamic stiffness of an air-spring. In: ISMA2004 international conference on noise & vibration engineering, pp 20–22
39. Pelz PF, Rösner J (2008) Schwingungsfluiddämpfung- und/oder -federung. Patent DE102008007566B4, 05 Feb 2008
40. Pelz PF, Holl M, Platzer M (2016) Analytical method towards an optimal energetic and economical wind-energy converter. *Energy* 94:344–351. <https://doi.org/10.1016/j.energy.2015.10.128>
41. Phadke MS (1989) *Quality engineering using robust design*. Prentice-Hall, London
42. Rams D, Klatt J (1995) Weniger, aber besser: Skizze für Phonokombination TP 1. Jo Klatt Design und Design Verlag, Hamburg
43. Rexer M, Brötz N, Pelz PF (2020) Much does not help much: 3D pareto front of safety, comfort and energy consumption for an active pneumatic suspension strut. In: 12th international fluid power conference, pp 37–42. <https://doi.org/10.25368/2020.91>

44. Saltelli A, Aleksankina K, Becker W, Fennell P, Ferretti F, Holst N, Li S, Wu Q (2019) Why so many published sensitivity analyses are false: a systematic review of sensitivity analysis practices. *Environ Model & Softw* 114:29–39. <https://doi.org/10.1016/j.envsoft.2019.01.012>
45. Scheitza M (2009) Konzeption eines flexiblen 3D-Servo-Pressensystems und repräsentative Basisanwendungen. Shaker
46. Sethi A, Sethi S (1990) Flexibility in manufacturing: a survey. *Int J Flex Manuf Syst* 2(4). <https://doi.org/10.1007/BF00186471>
47. Sinz J, Niessen B, Groche P (2018) Combined roller and plain bearings for forming machines: design methodology and validation. In: Congress of the German academic association for production technology. Springer, pp 126–135
48. Sivak M, Schoettle B (2015) Motion sickness in self-driving vehicles. Technical report, UMTRI-2015-12, Transportation Research Institute, University of Michigan. <http://hdl.handle.net/2027.42/111747>
49. Sivia DS, Skilling J (2006) Data analysis: a Bayesian tutorial, 2nd edn. Oxford University Press, Oxford
50. Stevin S (1582) Tafelen van interest. Christoffel Plantijn, Antwerpen. Neudruck
51. Sullivan LH (1922) The tall office building artistically considered. *West Archit* 1922(31):3–11
52. Vaughan D (2007) The Challenger launch decision: Risky technology, culture, and deviance at NASA. University of Chicago Press, Chicago. Reprint
53. Vergé A, Pötting P, Altherr LC, Ederer T, Pelz PF (2016) Lebensdauer als Optimierungsziel – Algorithmische Struktursynthese am Beispiel eines hydrostatischen Getriebes. *O+P – Ölhdraulik und Pneumatik* 60(1–2):114–121
54. Wegener K (2014) Forming presses (hydraulic, mechanical, servo). In: Laperrière L, Reinhart G (eds) *CIRP Encyclopedia of production engineering*. Springer, Berlin, pp 547–553. https://doi.org/10.1007/978-3-642-20617-7_16695

Open Access This book is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this book are included in the book's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the book's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

