TECHNISCHE UNIVERSITÄT MÜNCHEN

Lehrstuhl für Angewandte Mechanik

Dynamics of Variable Valve Trains and Extrapolation Methods for Time-Stepping Schemes

Dipl.-Ing. Univ. Robert Huber, M.Sc., M.Sc.

Vollständiger Abdruck der von der Fakultät für Maschinenwesen der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktor-Ingenieurs

genehmigten Dissertation.

Vorsitzender:

Univ.-Prof. Dr.-Ing. Horst Baier

Prüfer der Dissertation:

1. Univ.-Prof. Dr.-Ing. Dr.-Ing. habil. Heinz Ulbrich
2. Univ.-Prof. Dr.-Ing. Georg Wachtmeister

Die Dissertation wurde am 25.04.2012 bei der Technischen Universität München eingereicht und durch die Fakultät für Maschinenwesen am 30.11.2012 angenommen.

Contents

1 Introduction

Increasingly stricter limits for CO_2 emissions, rising energy costs, and growing concerns over energy security—especially the availability of fossil fuels—have led to the current trend towards electrification. Vehicle electrification varies from merely replacing inefficient belts and gear driven auxiliary units (such as pumps, cooling fans, or air conditioning) with efficient electric powered devices to entirely electric driven battery powered vehicles. Electrification can also combine electric propulsion with internal combustion engines in various parallel, series, power-split, and plug-in hybrid configurations to increase efficiency. Hybrid vehicles benefit by avoiding inefficient operating points such as partial load of the combustion engine and recuperating kinetic energy during deceleration. Generally, as vehicle power becomes more electric and fuel efficiency improves both cost and complexity increase.

Because the increasing trend towards electrification of automotive powertrains is accompanied by many unsolved problems such as high battery cost, limited electric range, uncertain reliability of electric components, and electric power's carbon footprint, as well as incomplete public awareness about the practicality of these vehicles, the internal combustion engine will continue to play a decisive role in the next few decades. The competition between combustion engine and electric drive will force innovations in both engine development and design. Reducing fuel consumption is the automotive industry's main task. Hence, the combustion engine is constantly in the focus of current developments towards efficient powertrains. Besides direct injection, one possible method of meeting ambitious reduction in fuel consumption and the emissions target is the use of variable valve trains.

Some conflicts of goals arise designing valve train systems due to the nature of fuel injection and gas exchange of current internal combustion engines. Engineers have to compromise among engine characteristics as fuel consumption, torque-progress, maximum power, idle stability, and emissions. Additional degrees of freedom are needed to meet the all these demands. Fully variable valve trains provide comprehensive means of adjustment in terms of variable valve timing and valve lift. The efficiency of the engine is improved in the operating range, but in return increasingly complex mechanical and control design must be handled. For optimisation and design of such kinds of complex systems, detailed and efficient simulation models are needed that cover different physical domains such as mechanics, hydraulics, electrodynamics and control.

Due to the multitude of different contacts within a valve train—namely coil collision of the valve spring, closing valves, contacts between cam and roller, frictional impacts at gearings or contacts within chain drives—it is obvious that within valve train simulations the main focus has to be the treatment of contact problems. A very

efficient concept to deal with contacts is to use methods of non-smooth mechanics in combination with time-stepping integration schemes. Hence, in the first part of this work a short introduction to non-smooth multi-body dynamics is given. In the second part of this work extrapolation methods are used to improve the well known time-stepping integration scheme. Also, parallel computing methods to speed up integration time are introduced.

The topic of the third part of this work is the variable valve train Audi valvelift system (AVS). A detailed simulation model of the entire timing assembly is derived, starting with the components. Specific component test rigs are designed to validate components including valve springs, hydraulic lash adjusters, electromagnetic actuators, and switch units. Lastly, the overall simulation model and its capabilities to optimise the system are shown.

Part I

Multi-Body Dynamics

2 Non-smooth Multi-Body Dynamics

In this chapter the governing equations describing the system dynamics and the theoretical framework for smooth and non-smooth multi-body systems are presented. Non-smooth mechanics concern mechanical situations with possible non-differentiable relationships such as unilateral contacts, dry friction, collisions, plasticity, damage, and phase transition without any regularisation process.
Non-smoothness could be seen as simplified mechanical modelling, because almost all dynamic systems in nature are smooth if the time scale is fine enough. Depending on the time scale of interest, non-smooth modelling in many cases leads to better approximations than smooth modelling. This is true especially against the background that approaches which try to reach nature exactly need detailed models and parameters that are often unknown or not available. Furthermore, such models require a large number of variables. Therefore, the mathematical formulation becomes non-smooth and regularising would only be an arithmetic trick without any physical justification.

The configuration of a set of interconnected bodies is identified by a set of variables called generalised coordinates $\boldsymbol{q}$ that completely define the location and orientation of each body. For interconnected systems these generalised coordinates, however, are not independent but restricted by constraints such as mechanical joints or contacts between adjacent bodies. They are formulated by algebraic equalities and inequalities. Depending on the type of constraint, the dynamics of such systems can be generally described in terms of ordinary differential equations (ODE), differential algebraic equations (DAE), or measure differential equations (MDE).

Unconstrained systems and systems subjected to bilateral (equality) constraints can be formulated as ODEs and DAEs respectively. An introductive depiction of DAE formulations for constrained systems is given in Section 2.1. Unilateral constraints, mathematically written as inequality expressions, yield to jumps within the state variables, generally the system velocities. The resulting dynamic is said to be non-smooth. Therefore, MDEs provide an applicative formulation. Section 2.2 gives a short introduction to non-smooth mechanics with special emphasis on the treatment of contact and impact problems with friction. The kinematics of contacts, as well as set-valued force laws and their appropriate formulation using convex analysis are presented. The methods presented in this work are not restricted to mechanical systems since they are also applicable to electrical networks [77, 144] and hydraulic systems [21]. The extension for hydraulic systems is shown in Section 2.3. The mathematical framework developed in this chapter will give the basis for numerical simulation methods and integration schemes presented in Chapter 3.

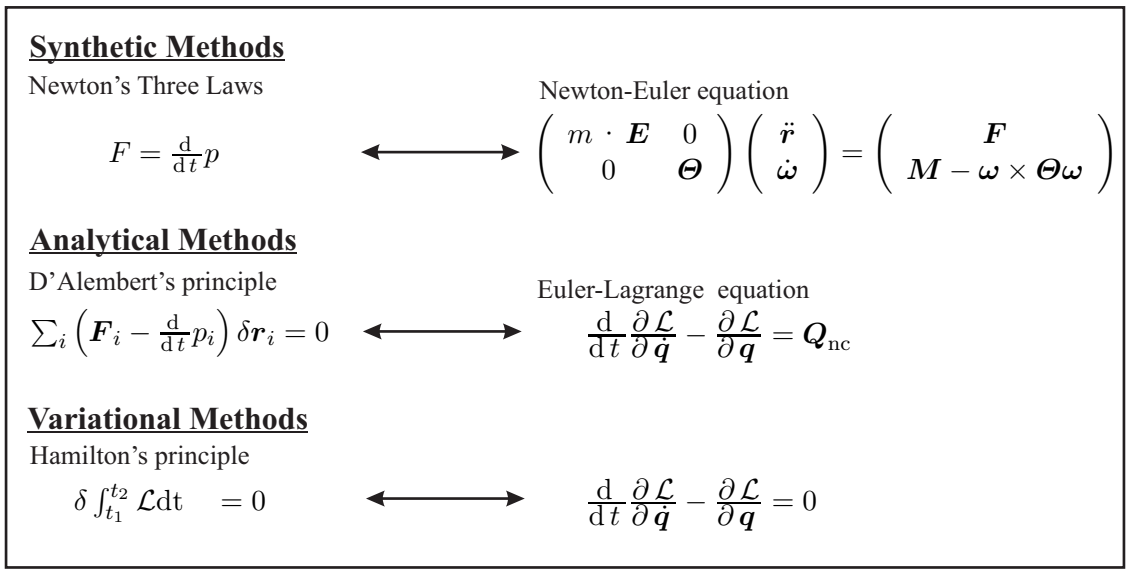


Figure 2.1: Overview: formulations in classical mechanics

2.1 Equations of Motion (ODE/DAE Formulation)

2.1.1 Theory of Classical Mechanics

The differential equations describing the dynamics of mechanical systems are derived using the classical theory of mechanics. In this section only a short overview is given. For further reading see e.g. [57,81,88]. To develop an axiomatic formulation of this theory some fundamental assumptions about time, space, and mass must be made. The most relevant assumptions are that space is complying with Euclidean geometry, time and space are absolute, mass is constant and in particular independent from velocity. Besides these assumptions Newton's three laws of motion are one possibility to form the axiomatic basis. Classical dynamics is a deductive theory based on postulates arising from experimental observations, and as a consequence it is only valid within the accuracy limits of the observations. As there is no evidence for time and space being absolute, these assumptions were abandoned by Einstein resulting in the theory of relativity. By omitting quantum mechanics and relativistic effects, speeds approaching the velocity of light and dimensions on an atomic or cosmic scale are eliminated from consideration. With this limitation, which in fact are no restrictions for most technical systems, classical mechanics is found to be in excellent agreement with observations over a broad range of system dimensions and velocities.

The theory of classical mechanics provides several equivalent formulations, which can all be used to describe the motion of a system in a mathematical model. Synthetic methods are based on Newton's second law relating force and momentum. In combination with Euler's concept of angular momentum conservation for rigid bodies, the Newton-Euler or projective Newton-Euler equations [197] directly provide the system dynamics in terms of second order ODE systems. Due to the vectorial

nature of Newtonian mechanics it may be more convenient in certain cases to use a second formulation referred to as analytical mechanics, which focuses on scalar energy functions. Built upon two basic postulates, the principle of virtual work by Bernoulli and the extension to dynamic systems of D'Alembert, Lagrangian or Hamiltonian mechanics allows set-up of the equations of motion by differentiating single scalar energy functions, namely the Lagrange $\mathcal{L}$ or Hamilton function $\mathcal{H}$ relating kinetic T and potential energy V. Historically, the study of ODEs has proceeded concurrently with the study of calculus of variation yielding variational principles in classical mechanics. These principles enable motion to be considered in a more metaphysical way: Variational approaches arise from the philosophical belief that there exists a compelling tendency towards rational simplicity in the physical universe. Hamilton's principle states that a conservative system always behaves such that the time integral of its Lagrangian function has a minimum or stationary value compared to neighbouring geometrically possible trajectories.

2.1.2 Unconstrained Systems

The above-presented concepts of classical mechanics are not a complete list, but provide an overview of the most relevant formulations (see Figure 2.1). All formulations of classical mechanics are equivalent and can be transformed into each other. The formulation one chooses to derive the equations of motion depends on the concrete system and personal preferences. Thus, orderly procedures for obtaining the equations of motion are available, yielding a set of differential equations (ODEs) for systems without constraints:

$$\boldsymbol{M}(\boldsymbol{q}) \cdot \ddot{\boldsymbol{q}} = \boldsymbol{h}(\dot{\boldsymbol{q}}, \boldsymbol{q}, t) \,. \tag{2.1}$$

The symmetric and positive definite mass matrix $\boldsymbol{M}$ depends on the generalised positions $\boldsymbol{q}$. The vector $\boldsymbol{h}$ denotes all smooth external, internal, and gyroscopic forces depending on $\boldsymbol{q}$, the generalised velocities $\dot{\boldsymbol{q}}$, and time t.

2.1.3 Constrained Systems

Equation (2.1) holds for unconstrained systems or systems described by a complete set of independent coordinates called minimal coordinates. As minimal coordinates specify only feasible system configurations, the constraints are implicitly fulfilled. For systems described by dependent coordinates, the equations of motion must be extended by a set of independent algebraic equations of constraints relating the coordinates. Bilateral constraints can be either *holonomic* or *non-holonomic* depending on whether generalised velocities are involved:

$$\boldsymbol{g}(\boldsymbol{q}, t) = 0 \qquad \text{holonomic} \tag{2.2}$$

$$\boldsymbol{\gamma}(\dot{\boldsymbol{q}}, \boldsymbol{q}, t) = 0 \qquad \text{non-holonomic} \,. \tag{2.3}$$

One possible way to treat holonomic systems is to introduce a new set of independent coordinates that satisfy all constraints, such as there is the same number of coordinates as degrees of freedom. As this approach is generally not possible for non-holonomic systems and not always desirable for holonomic systems (e.g. to simplify the model setup), the Lagrangian multiplier method can be used to extend the equations of motion by generalised force terms representing constraint forces.

Holonomic Bilateral Constraints

For holonomic bilateral systems the variational problem can be formulated in the presence of constraints by adding an additional term $\mathcal{L}_1$ comprising the independent Lagrangian multiplier $\boldsymbol{\lambda}(t)$ and the constraint function $\boldsymbol{g}(\boldsymbol{q},t)$ to the Lagrangian function:

$$\mathcal{L} = \mathcal{L}_0 + \mathcal{L}_1 = T - V + \boldsymbol{g}^T \cdot \boldsymbol{\lambda} \,. \tag{2.4}$$

The term $\mathcal{L}_1 = \boldsymbol{g}^T \cdot \boldsymbol{\lambda}$ can be interpreted as a kind of extra potential energy. As long as the system remains within the constraint surface, the constraint function $\boldsymbol{g}$ is equal to zero and thus the "extra" potential remains constant.
To obtain the equations of motion the conditional minimisation problem can be solved by evaluating the Euler-Lagrange equations for the augmented Lagrangian function (2.4):

$$\frac{\mathrm{d}}{\mathrm{d}\,t}\left(\frac{\partial\,\mathcal{L}}{\partial\,\dot{\boldsymbol{q}}}\right)^T - \left(\frac{\partial\,\mathcal{L}}{\partial\,\boldsymbol{q}}\right)^T - \boldsymbol{Q}_{\mathrm{nc}} = 0 \tag{2.5}$$

$$\underbrace{\frac{\mathrm{d}}{\mathrm{d}\,t}\left(\frac{\partial\,\mathcal{L}_0}{\partial\,\dot{\boldsymbol{q}}}\right)^T - \left(\frac{\partial\,\mathcal{L}_0}{\partial\,\boldsymbol{q}}\right)^T - \boldsymbol{Q}_{\mathrm{nc}}}_{\boldsymbol{M}\,\cdot\,\ddot{\boldsymbol{q}} - \boldsymbol{h}} + \underbrace{\frac{\mathrm{d}}{\mathrm{d}\,t}\left(\frac{\partial\,\boldsymbol{g}^T\boldsymbol{\lambda}}{\partial\,\dot{\boldsymbol{q}}}\right)^T}_{0} - \underbrace{\left(\frac{\partial\,\boldsymbol{g}}{\partial\,\boldsymbol{q}}\right)^T\boldsymbol{\lambda}}_{\boldsymbol{W}\,\cdot\,\boldsymbol{\lambda}} = 0 \tag{2.6}$$

yielding a set of differential equations for the system dynamics and algebraic equations expressing the constraints:

$$\boldsymbol{M}\,(\boldsymbol{q}) \cdot \ddot{\boldsymbol{q}} = \boldsymbol{h}\,(\dot{\boldsymbol{q}},\boldsymbol{q},t) + \boldsymbol{W}\,(\boldsymbol{q}) \cdot \boldsymbol{\lambda} \tag{2.7a}$$

$$\boldsymbol{g}\,(\boldsymbol{q},t) = 0\,. \tag{2.7b}$$

The Lagrangian multiplier $\boldsymbol{\lambda}$ can be interpreted as constraint forces. The gradient $\boldsymbol{W}^T = \frac{\partial\,\boldsymbol{g}}{\partial\,\boldsymbol{q}}$ is orthogonal to the manifold $\boldsymbol{g} = 0$ of the constraint surface and therefore represents the direction of the constraint forces. Hence, the constraint matrix $\boldsymbol{W}$ is a transformation from the space of constraints to the configuration space. The constraint forces are not known directly and are included as additional unknown variables in the equations of motion. Therefore, the constraint equation (2.7b) must be added to the equations of motion to yield a DAE system.

Non-Holonomic Bilateral Constraints

Up to this point only constraints that are expressed by position coordinates have been considered. However, important cases exist—such as constraints considering Coulomb friction like a disk rolling on a surface without sliding—that can be described only using constraints involving velocities, the so called non-integrable or non-holonomic constraints. It seems obvious to add non-holonomic constraints to the Lagrangian function in the same way as in the holonomic case. But it turns out that the Euler-Lagrangian equations do not result in correct equations of motion. The main reason for this is that the velocities are not varied independently from the position parameters. Thus, the generalised forces representing the non-holonomic constraint forces have to be introduced using the principle of virtual work.
According to Pfeiffer [160] non-holonomic constraints in mechanics are linear in velocities:

$$\boldsymbol{\gamma}\left(\dot{\boldsymbol{q}}, \boldsymbol{q}, t\right) = \boldsymbol{W}^T\left(\boldsymbol{q}, t\right) \cdot \dot{\boldsymbol{q}} + \boldsymbol{w}\left(\boldsymbol{q}, t\right) = 0\,. \tag{2.8}$$

Note that with $\boldsymbol{W}^T = \frac{\partial \boldsymbol{g}}{\partial \boldsymbol{q}}$ and $\boldsymbol{w} = \frac{\partial \boldsymbol{g}}{\partial t}$ holonomic constraints can be transformed into the form of Eq. (2.8) by differentiating and using the velocity level $\dot{\boldsymbol{g}}$ as constraint equation (cf. Eq. (2.16)). Thus, subsequent formulations will be found to be applicable for the holonomic as well as for the non-holonomic case.
The differential form of the non-holonomic constraint Eq. (2.8) is written as:

$$\boldsymbol{W}^T \mathrm{d}\boldsymbol{q} + \boldsymbol{w}\mathrm{d}t = 0\,. \tag{2.9}$$

At any instant of time the variation $\delta\boldsymbol{q}$ of the generalised coordinates in a virtual displacement must meet the condition

$$\boldsymbol{W}^T \delta\boldsymbol{q} = 0 \tag{2.10}$$

as the variation δt in time is equal to zero. The virtual displacement is consistent with the instantaneous constraints of the system and therefore no work is done by the corresponding generalised (constraint) forces $\boldsymbol{Q}_\mathrm{c}$:

$$\boldsymbol{Q}_\mathrm{c}^T \delta\boldsymbol{q} = 0\,. \tag{2.11}$$

First, a single constraint is considered. The constraint matrix $\boldsymbol{W}$ only consists of a single column vector $\boldsymbol{W}_i$. Now Eq. (2.10) is multiplied with a scalar factor λ_i, known as Lagrangian multiplier:

$$\boldsymbol{W}_i^T \delta\boldsymbol{q} \lambda_i = 0\,. \tag{2.12}$$

Next, Eq. (2.12) is subtracted from Eq. (2.11) yielding:

$$\left(\boldsymbol{Q}_{\mathrm{c},i}^T - \boldsymbol{W}_i^T \lambda_i\right) \delta\boldsymbol{q} = 0\,. \tag{2.13}$$

So far, λ_i has been considered as arbitrary, whereas the variations $\delta\boldsymbol{q}$ satisfy the constraints. However, it is possible to choose λ_i such that the coefficient of $\delta\boldsymbol{q}$

vanishes. Then Eq. (2.13) will apply for any value of $\delta\boldsymbol{q}$. They can be chosen independently if the constraint force $\boldsymbol{Q}_{\mathrm{c},i} = \boldsymbol{W}_i\lambda_i$ is taken into account.
For more than one constraint each column of the matrix $\boldsymbol{W}$ represents a single constraint and $\boldsymbol{\lambda}$ is the vector of all associated Lagrangian multipliers. The resulting reaction of all constraints is obtained by the superimposition principle:

$$\boldsymbol{Q}_{\mathrm{c}} = \sum_i \boldsymbol{Q}_{\mathrm{c},i} = \boldsymbol{W}\boldsymbol{\lambda}\,, \tag{2.14}$$

resulting in equations of motion with an analogue structure as for the holonomic case:

$$\boldsymbol{M}\,(\boldsymbol{q}) \;\cdot\; \ddot{\boldsymbol{q}} = \boldsymbol{h}\,(\dot{\boldsymbol{q}},\boldsymbol{q},t) + \boldsymbol{W}\,(\boldsymbol{q}) \;\cdot\; \boldsymbol{\lambda} \tag{2.15a}$$

$$\boldsymbol{\gamma}\,(\dot{\boldsymbol{q}},\boldsymbol{q},t) = 0\,. \tag{2.15b}$$

The only difference between the time derivative of holonomic constraints $\dot{\boldsymbol{g}}$ and non-holonomic constraints $\boldsymbol{\gamma}$ is that $\dot{\boldsymbol{g}}$ can be integrated to $\boldsymbol{g}$ whereas $\boldsymbol{\gamma}$ is not integrable. In most cases this distinction is not relevant. Unless otherwise indicated, $\dot{\boldsymbol{g}}$ is used in the following sections for both holonomic constraints on the velocity level, as well as non-holonomic constraints. For non-holonomic constraints $\boldsymbol{\gamma}$ is only used for emphasizing that the constraint is not integrable.

Differential Algebraic Equation Formulation

Systems such as Eq. (2.7) and Eq. (2.15) are called differential algebraic equations (DAE). The differentiation index plays an important role in the numerical treatment of such DAEs. It is the minimal number of analytical differentiations that is needed to transform a DAE system into an explicit ODE system for all unknowns, namely generalised coordinates $\boldsymbol{q}$ and Lagrangian multipliers $\boldsymbol{\lambda}$. Hence, the index is a non-negative integer that provides useful information about the mathematical structure of the DAE, and is in addition a measure of the singularity. In general, the higher the index the more difficulties can be expected for its numerical solution. Stability and convergence cannot be guaranteed [159]. Errors and perturbations due to finite precision arithmetic pollute the numerical solution of high index DAEs [10,48]. This pollution can cause disastrous effects for small values of the time step size. State variables and Lagrangian multipliers are affected by increasing errors as the time step size decreases [158]. A short overview of numerical methods for the solution of DAEs is given by Cash [30].
Mechanical systems with constraints on the position level as in Eq. (2.7) are of index 3. As only a few straight forward integration algorithms exist for index 3 formulation (e.g. the implicit Runge-Kutta method RADAU5 [86, 87] from Hairer and Wanner, half-explicit Runge-Kutta methods [156] or Botassos's preconditioning approach [22]), index reduction by differentiating the constraint equation $\boldsymbol{g} = 0$ is an usual technique for such systems:

$$\boldsymbol{g} = 0 \qquad \Leftrightarrow \qquad \dot{\boldsymbol{g}} = \boldsymbol{W}^T \;\cdot\; \dot{\boldsymbol{q}} + \boldsymbol{w} = 0 \quad \vee \quad \boldsymbol{g}(\boldsymbol{q}_0,t_0) = 0\,. \tag{2.16}$$

Constraints on the position level can be described on the velocity level if the initial conditions are fulfilled. Replacing constraints on the position level (Eq. (2.7b)) by constraints on the velocity level yield the equations of motion in an index 2 formulation. Similarly, the index can be reduced to index 1 by describing the constraints on the acceleration level:

$$\begin{aligned} \boldsymbol{g} = 0 \qquad \Leftrightarrow \qquad \ddot{\boldsymbol{g}} = \boldsymbol{W}^T \cdot \ddot{\boldsymbol{q}} + \tilde{\boldsymbol{w}} = 0 \quad &\vee \quad \boldsymbol{g}(\boldsymbol{q}_0, t_0) = 0 \\ &\vee \quad \boldsymbol{\gamma}(\dot{\boldsymbol{q}}_0, \boldsymbol{q}_0, t_0) = 0\,. \end{aligned} \tag{2.17}$$

Substituting $\ddot{\boldsymbol{q}}$ in the constraints on the acceleration level by the equations of motion, the constraint force can be derived explicitly:

$$\boldsymbol{\lambda} = \left(\boldsymbol{W}^T \boldsymbol{M}^{-1} \boldsymbol{W}\right)^{-1} \left(-\tilde{\boldsymbol{w}} - \boldsymbol{W}^T \boldsymbol{M}^{-1} \boldsymbol{h}\right)\,. \tag{2.18}$$

Differentiating Eq. (2.18) yields an ODE in the unknown variables $\boldsymbol{\lambda}$ and $\boldsymbol{q}$. Thus, the equations of motion for multi-body systems (Eq. (2.7)) are index 3 DAEs. Non-holonomic constraints $\boldsymbol{\gamma}$ can be treated equally to holonomic constraints $\dot{\boldsymbol{g}}$ on the velocity level.
The formulations with index 1 to 3 are mathematically equivalent. However, if systems with index less than 3 are integrated numerically, the constraint $\boldsymbol{g}$ on position level will no longer be satisfied. Since the derivatives of the constraint $\boldsymbol{g}$ are not met exactly, but only within the numerical accuracy, the error on the position level increases with time. To minimise this numerical drift, projection or stabilisation methods like Gear-Gupta-Leimkuler [75] or Baumgarte [15] stabilisation are applied.

2.1.4 Kinematic Equations

Since the angular velocities in the spatial case are not equal to the time derivative of the angular position parameter (e.g. Euler or Cardan angles or Euler parameter), it may be advantageous to introduce generalised velocities $\boldsymbol{u}$ that can be directly interpreted as angular velocities. Thus, kinematic equations have to be regarded by the linear relationship:

$$\dot{\boldsymbol{q}} = \boldsymbol{\Upsilon} \cdot \boldsymbol{u}\,. \tag{2.19}$$

The block-diagonal transformation matrix $\boldsymbol{\Upsilon} = \boldsymbol{\Upsilon}(\boldsymbol{q})$ is not necessarily square (for example if Euler parameters are used). The blocks are composed of the degrees of freedom for the individual objects e.g. rotational and translational degrees of freedom for rigid bodies. Using Eq. (2.19) the equations of motion for constrained systems can be rewritten as:

$$\overline{\boldsymbol{M}} \cdot \dot{\boldsymbol{u}} = \bar{\boldsymbol{h}} + \overline{\boldsymbol{W}} \cdot \boldsymbol{\lambda}\,, \tag{2.20}$$

with the transformation formulas:

$$\overline{\boldsymbol{M}} = \boldsymbol{\Upsilon}^T \boldsymbol{M} \boldsymbol{\Upsilon} \tag{2.21a}$$

$$\overline{\boldsymbol{h}} = \boldsymbol{\Upsilon}^T \boldsymbol{h} - \boldsymbol{\Upsilon}^T \boldsymbol{M} \dot{\boldsymbol{\Upsilon}} \boldsymbol{u} \tag{2.21b}$$

$$\overline{\boldsymbol{W}} = \boldsymbol{\Upsilon}^T \boldsymbol{W} = \left(\frac{\partial \dot{\boldsymbol{g}}}{\partial \boldsymbol{u}} \right)^T . \tag{2.21c}$$

Setting up the equations of motion using $[\boldsymbol{q}; \dot{\boldsymbol{q}}]$ or $[\boldsymbol{q}; \boldsymbol{u}]$ as state variables yields equations of the same structure that relate state variables, mass matrix, force vector, and constraint reactions. As it is apparent from the context which formulation is used, in the following chapters we will not distinguish the involved matrices and vectors by any labelling such as an overbar. One must be aware that mass matrix, constraint matrix and force vector for both formulations are identical only in the case that the transformation matrix $\boldsymbol{\Upsilon}$ is an identity matrix (cf. Eq. 2.21).
In the next chapters the following formulation of the equations of motion for systems with holonomic and non-holonomic constraints is used. Since the structure of Eq. (2.8) and (2.16) is similar, the constraint equations $\dot{\boldsymbol{g}}$ contain non-holonomic as well as holonomic constraints on the velocity level:

$$\dot{\boldsymbol{q}} = \boldsymbol{\Upsilon} \cdot \boldsymbol{u} \tag{2.22a}$$

$$\boldsymbol{M} \cdot \dot{\boldsymbol{u}} = \boldsymbol{h} + \boldsymbol{W} \cdot \boldsymbol{\lambda} \tag{2.22b}$$

$$\dot{\boldsymbol{g}} = \boldsymbol{W}^T \cdot \boldsymbol{u} + \boldsymbol{w} = 0 . \tag{2.22c}$$

2.2 Equations of Motion for Non-Smooth Systems

The equations of motion of the previous section express non-impulsive dynamics of a constrained system at acceleration level under the assumption of state variables $\boldsymbol{u}$ and $\boldsymbol{q}$ being continuously differentiable. Hence, the acceleration $\dot{\boldsymbol{u}}$ is continuous for all times and the corresponding systems are called smooth. According to Peano's theorem [157] a solution exists for such systems.
In many applications physical phenomena such as friction, impacts, or backlash result in mathematical models with discontinuities making methods of non-smooth mechanics necessary; see among others Brogliato [25, 26] for a general overview and Pfeiffer [160], Pfeiffer and Glocker [161], and Glocker [76] for elaborations focused on mechanical systems. The term non-smooth refers to systems whose trajectories are not overall differentiable. Since the existence theorem of Peano is not valid in this case, the corresponding initial value problem may have no solution. Standard ODE theory is not applicable and special theory and solution strategies are necessary.

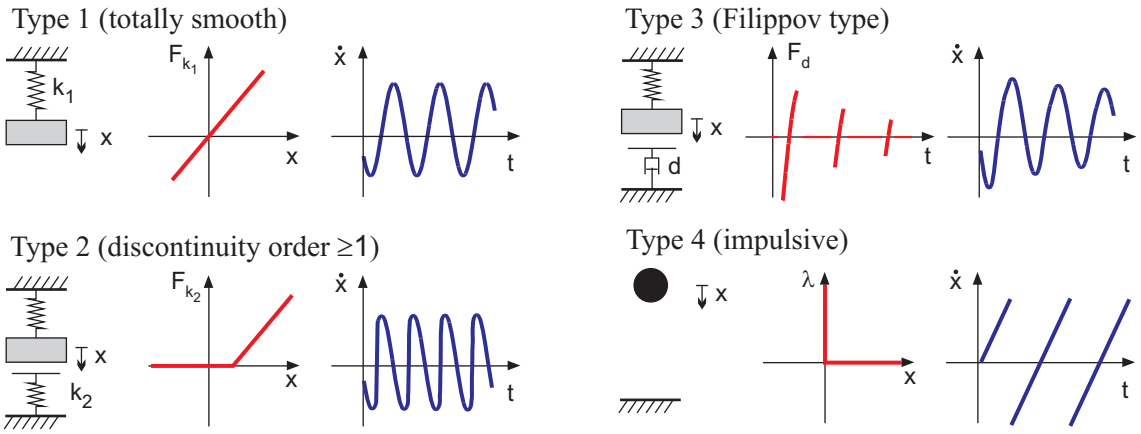


Figure 2.2: Classification of smooth and non-smooth systems

2.2.1 Classification of Non-Smooth Systems

In literature the term non-smooth is often used for various systems without stating explicitly which properties of the system are meant. In this work dynamic systems $\dot{\boldsymbol{z}} = \boldsymbol{f}(\boldsymbol{z},t)$ are classified by continuity properties of the right-hand side function $\boldsymbol{f}(\boldsymbol{z},t)$ and divided into four types according to their degree of smoothness (see Figure 2.2):

1. The vector field $\boldsymbol{f}$ is totally smooth or in other words infinitely differentiable. Hence, standard ODE theory is applicable and the existence of a solution is guaranteed.

2. The function $\boldsymbol{f}$ itself is continuous and has at least one discontinuity in the partial derivative of order $q \geq 1$ and has continuous derivatives of order zero to $q-1$. In most issues ODE theory systems belonging to this type can be treated equally to systems of the first type. However, attention has to be paid to discontinuous gradients in numerical solving methods and stability analyses; e.g. the time integration order is restricted by the order of discontinuity [29]. An example for this type of system is a single mass oscillator one-sided supported by a linear spring.

3. The right-hand side function is discontinuous, but the states are continuous (Filippov systems). In order to guarantee the existence of a solution additional assumptions have to be made e.g. no spontaneous or infinite jumps are allowed. Examples for this type of non-smooth system are a single mass oscillator with one-sided viscoelastic support or dry friction.

4. Systems of this type of non-smoothness are characterised by discontinuities in the state variables e.g. impact systems with velocity reversals due to impacts. To obtain a solution for such impulsive systems additional impact laws are necessary.

Type 3 and 4 systems are called non-smooth. The source of the non-smoothness of Filippov systems are discontinuous but bounded functions that contribute to

the right-hand side e.g. discrete controllers or force switching elements. Classical examples for type 4 systems are inequality constraints e.g. due to set-valued force laws. With changing constraints the dimension of the differential equation is not necessarily constant resulting in a time-variant system topology. The last example in Figure 2.2 is a ball impacting a rigid floor. The contact force at the instant of impact can be described mathematically as a Dirac pulse. In the next section the measure differential equation (MDE) is introduced for a general description of such impulsive systems. For a more detailed classification of non-smooth systems consult Acary and Brogliato [3].

2.2.2 Unilateral Constraints and Impacts

In Section 2.1 bilateral or equality constraints of the form $\boldsymbol{g}_B = 0$ (index 3 formulation) or $\dot{\boldsymbol{g}}_B = 0$ (index 2 formulation) have been considered. Unilateral constraints are expressed by an inequality equation:

$$\boldsymbol{g}_U(\boldsymbol{q}, t) \geq 0\,. \tag{2.23}$$

Unilateral constraints are, for example, used to describe normal contact conditions. Equations for unilateral constraints are derived by sorting between smooth transitions and discrete impacts. The variational problem in the presence of unilateral constraints is formulated in the same way as for bilateral constraints by adding the term $\mathcal{L}_1 = \boldsymbol{g}_U^T \cdot \boldsymbol{\lambda}_U$ with the positive Lagrangian multiplier $\boldsymbol{\lambda}_U \geq 0$ to the Lagrangian function:

$$\mathcal{L} = T - V + \boldsymbol{g}_U^T \boldsymbol{\lambda}_U\,. \tag{2.24}$$

As the "extra" potential $\mathcal{L}_1$ should have no effect if the constraint is fulfilled, the equation $\boldsymbol{g}_U^T \boldsymbol{\lambda}_U = 0$ must be met. Evaluating the Euler-Lagrange equations for Eq. (2.24) yields:

$$\boldsymbol{M}\dot{\boldsymbol{u}} = \boldsymbol{h} + \boldsymbol{W}_U \boldsymbol{\lambda}_U \tag{2.25a}$$

$$\boldsymbol{g}_U \geq 0, \quad \boldsymbol{\lambda}_U \geq 0, \quad \boldsymbol{g}_U^T \boldsymbol{\lambda}_U = 0\,. \tag{2.25b}$$

Equation (2.25a) describes only the smooth part of the motion. Additional impact laws are needed for discontinuities such as closing contacts. These laws must at least ensure post impact validity of the constraint equations. The complementary condition ($\boldsymbol{g}_U \geq 0$, $\boldsymbol{\lambda}_U \geq 0$, $\boldsymbol{g}_U^T \boldsymbol{\lambda}_U = 0$), which arises from a variational problem, can also be interpreted as description of a mechanical contact; only active or closed contacts ($v g_N = v g_U = 0$) imply positive contact forces $\boldsymbol{\lambda}_N = \boldsymbol{\lambda}_U$ in normal direction, whereas for inactive contacts ($\boldsymbol{g}_N = \boldsymbol{g}_U > 0$) the contact force is equal to zero.
Hence, active contacts can be treated in the same way as bilateral contacts as long as the contact force is positive. Inactive contacts have no influence on the system dynamics. Additional equations are necessary for transitions. Corresponding equations can be derived for unilateral constraints on the velocity level.

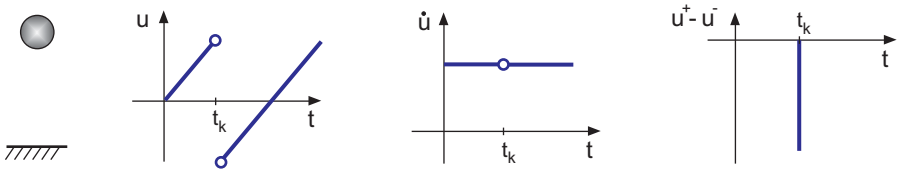


Figure 2.3: Smooth and non-smooth velocity components

2.2.3 Measure Differential Equations

The equations of motion in its presented form cannot represent discontinuous motion because at discontinuities t_k such as impacts, the acceleration $\dot{\boldsymbol{u}}$ is not defined. In Figure 2.3 the continuous and discontinuous parts of the velocity of an elastic impact are shown. Almost everywhere the derivative $\dot{\boldsymbol{u}}$ of $\boldsymbol{u}$ is equal to the gravitational acceleration, except for a single instant infinite peak at time t_k.
One possibility is to distinguish between continuous and discontinuous motion. Therefore, the set $\mathcal{M}_S$ containing all discontinuities t_k is defined. At $t = t_k$ special impact laws are applied, whereas between t_k and t_{k+1} the motion is described with standard ODE/DAE formulations. This approach is the basic concept for integration schemes with event detection.
Another suitable and more efficient way to describe impulsive systems is to use measure differential equations (Eq. (2.26)) introduced by Moreau [145]. This description leads in a natural way to a special class of numerical integration schemes, the so called time-stepping schemes.

$$\boldsymbol{M}\mathrm{d}\boldsymbol{u} = \boldsymbol{h}\mathrm{d}t + \boldsymbol{W}\mathrm{d}\boldsymbol{\Lambda} \tag{2.26}$$

At the discontinuities the standard notation of the derivative is not defined. Hence, to allow for finite jumps within the velocities a measure for the velocities is introduced:

$$\mathrm{d}\boldsymbol{u} = \dot{\boldsymbol{u}}\mathrm{d}t + \left(\boldsymbol{u}^{+} - \boldsymbol{u}^{-}\right)\mathrm{d}\eta\,. \tag{2.27}$$

It is split in a Lebesgue integrable term $\dot{\boldsymbol{u}}\mathrm{d}t$ for the continuous part and a purely atomic part $(\boldsymbol{u}^{+} - \boldsymbol{u}^{-})\,\mathrm{d}\eta$ at the discontinuities. The left and right limits of the velocities at $t_k \in \mathcal{M}_S$ are denoted by $\boldsymbol{u}^{+}$ and $\boldsymbol{u}^{-}$. For technical systems we can assume that these limits always exist and are always bounded. Pre- and post-impact velocities are weighted by the sum of the Dirac point measure at t_k as defined in Eq. (2.28). An arbitrary measurable map is denoted by f and $\mathcal{T}$ is a time interval.

$$\mathrm{d}\eta = \sum_{\mathcal{M}_S} \mathrm{d}\delta_k \qquad \mathrm{d}\delta_k = \mathrm{d}\delta\left(t - t_k\right) \tag{2.28a}$$

$$\int_{\mathcal{T}} f(t)\mathrm{d}\delta_k = \begin{cases} f(t_k) & t_k \in \mathcal{T} \\ 0 & t \notin \mathcal{T} \end{cases} \tag{2.28b}$$

Since $\boldsymbol{u}^+$ and $\boldsymbol{u}^-$ are still bounded and the impacts are assumed to be discrete, the position parameter $\boldsymbol{q}$ remains constant during impacts with infinite length. As a consequence the position parameter and the depending matrices $\boldsymbol{M}$ and $\boldsymbol{W}$ are $\mathcal{C}^0$ continuous.
The reaction measure $\boldsymbol{W}\mathrm{d}\Lambda$ is composed by the generalised direction of the non-smooth constraint forces and the impulse magnitude $\mathrm{d}\boldsymbol{\Lambda}$. The reactions are also split into a Lebesgue measurable part $\boldsymbol{\lambda}\mathrm{d}t$ containing smooth forces and an atomic part $\boldsymbol{\Lambda}\mathrm{d}\eta$ for discrete impulses for example due to collisions or impacts:

$$\mathrm{d}\boldsymbol{\Lambda} = \boldsymbol{\lambda}\mathrm{d}t + \boldsymbol{\Lambda}\mathrm{d}\eta\,. \tag{2.29}$$

Integration of Eq. (2.26) under consideration of the Dirac measure Eq. (2.28) yields the classical equations of motion for continuous periods:

$$\boldsymbol{M}\dot{\boldsymbol{u}} = \boldsymbol{h} + \boldsymbol{W}\lambda \qquad \forall t \notin \mathcal{M}_S\,, \tag{2.30}$$

as well as for discrete impulsive points:

$$\boldsymbol{M}_k\left(\boldsymbol{u}_k^+ - \boldsymbol{u}_k^-\right) = \boldsymbol{W}_k\Lambda_k \qquad \forall t \in \mathcal{M}_S\,. \tag{2.31}$$

Both equations contain a priori unknown reactions $\boldsymbol{\lambda}$ and $\boldsymbol{\Lambda}$, respectively. Hence, additional constraint and impact laws connecting reactions and kinematics are necessary for the evaluation of Eq. (2.26).

2.2.4 Kinematics of Constraints

For the description of contacts between two bodies (index i = 1,2), the contact geometry needs to be analysed no matter which contact model is used. The potential contact points $\boldsymbol{r}_{\mathrm{C}_i}$ on the contours need to be determined, as well as the gap distance g_N and the relative velocities $\dot{g}_N$ and $\dot{\boldsymbol{g}}_T$. This information is used for the evaluation of the force laws, as well as for the calculation of the generalised force directions $\boldsymbol{W}$.

Contours are described by a position vector $\boldsymbol{r}_{\mathrm{C}_i} = \boldsymbol{r}_{\mathrm{C}_i}(\boldsymbol{q},\boldsymbol{s}_i)$, the inward pointing contour normal $\boldsymbol{n}_i = \boldsymbol{n}_i(\boldsymbol{q},\boldsymbol{s}_i)$ and the associated tangents $\boldsymbol{T}_i = [\boldsymbol{t}_i(\boldsymbol{q},\boldsymbol{s}_i),\boldsymbol{b}_i(\boldsymbol{q},\boldsymbol{s}_i)]$ including the bi-normal $\boldsymbol{b}_i$, all depending on the generalised position of the associated body and on the contour parameters $\boldsymbol{s}_i$. To avoid additional impacts due to the geometric parametrisation of the contour, the parametrisation must comprise parametric continuity of order 2, i.e. the first and second derivatives with respect to s are continuous. Hence, cubic splines or cubic B-splines are often used to describe non-primitive curves.
Potential contact points can be identified by the necessary contact condition, that the difference vector

$$\boldsymbol{r}_{\mathrm{D}} = \boldsymbol{r}_{\mathrm{C}_2} - \boldsymbol{r}_{\mathrm{C}_1} \tag{2.32}$$

between potential contact points on both bodies is perpendicular to both tangents:

$$\begin{aligned} \boldsymbol{r}_{\mathrm{D}}^{T}(\boldsymbol{s}_1, \boldsymbol{s}_2)\, \boldsymbol{T}_1(\boldsymbol{s}_i) &= 0 \\ \boldsymbol{r}_{\mathrm{D}}^{T}(\boldsymbol{s}_1, \boldsymbol{s}_2)\, \boldsymbol{T}_2(\boldsymbol{s}_i) &= 0\,. \end{aligned} \tag{2.33}$$

Depending on the structure Eq. (2.33) is solved analytically for geometric primitive contour pairings or numerically using root-finding algorithms (e.g. Newton's method, Bisection methods, or Regula Falsi) to get a set of potential contact parameters. Selecting the solution with minimal normal distance g_N:

$$g_N = \boldsymbol{r}_{\mathrm{D}}^{T}\, \boldsymbol{n}_2 \tag{2.34}$$

the relative normal and tangential velocities are calculated by the projection equations:

$$\begin{aligned} \dot{g}_N &= \left[\dot{\boldsymbol{r}}_{\mathrm{C}_2} - \dot{\boldsymbol{r}}_{\mathrm{C}_1}\right]^{T}\, \boldsymbol{n}_2 \\ \dot{\boldsymbol{g}}_T^{T} &= \left[\dot{\boldsymbol{r}}_{\mathrm{C}_2} - \dot{\boldsymbol{r}}_{\mathrm{C}_1}\right]^{T}\, \boldsymbol{T}\,. \end{aligned} \tag{2.35}$$

Based on the contact points the JACOBIAN $\boldsymbol{J}_{\mathrm{C}_i} = (\partial\, \dot{\boldsymbol{r}}_{C_i} / \partial\, \boldsymbol{u}\,)^T$ between the system-space and the generalised coordinates of the two bodies are developed. The generalised force directions are written as:

$$\begin{aligned} \boldsymbol{W}_N^{T} &= \boldsymbol{n}_1^{T} \boldsymbol{J}_{\mathrm{C}_1} + \boldsymbol{n}_2^{T} \boldsymbol{J}_{\mathrm{C}_2} \\ \boldsymbol{W}_T^{T} &= \boldsymbol{T}_1^{T} \boldsymbol{J}_{\mathrm{C}_1} + \boldsymbol{T}_2^{T} \boldsymbol{J}_{\mathrm{C}_2}\,. \end{aligned} \tag{2.36}$$

A detailed description of contact kinematics can be found in [76]. The calculation of a planar contact pairing, namely a circle and an arbitrary convex hyper-surface, is shown in section 6.3.3.

2.2.5 Force Laws

Solving the equations of motion by calculating the accelerations $\dot{\boldsymbol{u}}$ and post-impact velocities $\boldsymbol{u}_k^{+}$ requires knowledge of the unknown contact reactions $\boldsymbol{\lambda}$ and $\boldsymbol{\Lambda}$ respectively at the discontinuities. To establish different force laws, a decomposition of the contact force $\boldsymbol{\lambda}_C = \boldsymbol{n}\, \lambda_N + \boldsymbol{T}\, \boldsymbol{\lambda}_T$ in a component normal to the contact plane and tangential components is used. A bilateral or unilateral constraint may correspond to $\lambda_N \in \{\lambda_B, \lambda_U\}$. The tangential forces due to Coulomb's friction law are $\boldsymbol{\lambda}_T \in \mathbb{R}^2$. The decomposition for discrete impulses is carried out analogously.

Single-Valued Force Laws

A commonly used approach to model a unilateral contact considering elasticity is to use spring and damper elements to minimize penetration. Using linear elements

with spring stiffness c and damping coefficient d the force law is written as:

$$\lambda_N = \begin{cases} -c\, g_N \, - \, d\, \dot{g}_N & \text{if } g_N < 0 \\ 0 & else \ . \end{cases} \tag{2.37}$$

More sophisticated constitutive laws are obtained by using non-linear correlations. For example, according to HERTZ [100], exponential force laws for $g_N < 0$ [114] are used to apply contact stiffness. To avoid discontinuities in λ_N at the beginning of the penetration, d is set to zero for $g_N = 0$ and is increased with increasing penetration up to a maximum value. For bilateral contacts equivalent force laws can be stated leading to linear or non-linear constitutive laws that may be interpreted as spring-damper elements. These regularised or flexible force laws are illustrated in Figure 2.4.

For considering friction in a bilateral or closed unilateral contact, Coulomb's law with relative tangential velocity $\dot{\boldsymbol{g}}_T \in \mathbb{R}^2$ and friction coefficient μ is used:

$$\dot{\boldsymbol{g}}_T \, = \, \mathbf{0} \quad \Rightarrow |\boldsymbol{\lambda}_T| \, \leq \, \mu |\boldsymbol{\lambda}_N| \tag{2.38a}$$

$$\dot{\boldsymbol{g}}_T \, \neq \, \mathbf{0} \quad \Rightarrow \quad \boldsymbol{\lambda}_T \, = \, -\frac{\dot{\boldsymbol{g}}_T}{|\dot{\boldsymbol{g}}_T|} \mu |\boldsymbol{\lambda}_N| \, . \tag{2.38b}$$

To use this discontinuous friction model in smooth systems, integrable by standard ODE solvers, the discontinuities must be regularised as illustrated in Figure 2.4 for planar friction.

Set-Valued Force Laws

In the previous section contacts were modelled using spring and damper elements. Due to the high stiffness of the contacts this approach results in stiff differential equations with unintentional high eigenfrequencies and high computing times, as well as uncertainties in the parameters for contact stiffness and damping. Therefore, another possibility to model contacts, namely the rigid approach, is discussed. For the rigid approach the contact is described by a discrete point model where the stiffness in the contact zone is assumed to be infinite, resulting in a constraint (cp. Section 2.2.2). Mathematically this means that an explicit functional relation with respect to the state variables no longer exists (cp. Figure 2.4. The contact reactions are described by set-valued laws.

Contact laws

For unilateral contacts the assumption of no penetration can be expressed by the SIGNORINI-FICHERA-condition:

$$g_N \, \geq \, 0, \quad \lambda_N \, \geq \, 0, \quad g_N \lambda_N \, = \, 0 \, . \tag{2.39}$$

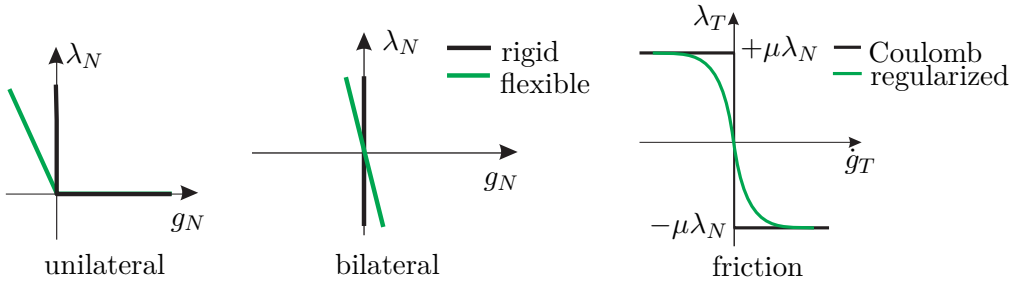


Figure 2.4: Set-valued force laws: uni- and bilateral contact and friction

In terms of set-valued force laws a bilateral contact implies a bilateral constraint of the form:

$$g_N = 0, \quad \lambda_N \in \mathbb{R}\,. \tag{2.40}$$

Using set-valued force laws to describe contacts with friction, Coulomb's law can be applied directly without any regularisation. Therefore, real sticking and transitions between stick and slip are modelled in a physically correct way.
Due to their mathematical nature set-valued force laws lead to jumps in velocities and therefore cannot be integrated with standard ODE solvers. Special integration schemes and solvers are needed for calculating the reaction forces $\boldsymbol{\lambda}$ and $\boldsymbol{\Lambda}$.

Impact laws

For impact formulations, the contact laws are transformed to the velocity level yielding for bilateral constraints:

$$\dot{g}_N^+ = 0, \quad \Lambda_N \in \mathbb{R}\,, \tag{2.41}$$

and for unilateral constraints:

$$\dot{g}_N^+ \geq 0, \quad \Lambda_N \geq 0, \quad \dot{g}_N^+ \Lambda_N = 0 \qquad \text{if } g_N = 0\,. \tag{2.42}$$

For the time discretization used within this study (cp. Section 3) Eq. (2.42) represents an inelastic impact as shown by Funk [74]. To consider elastic impacts described by Newton's impact, law the post-impact velocity $\dot{g}_N^+$ is replaced by a kind of mean velocity calculated from the sum of pre- and post-impact velocities weighted with the restitution coefficient ε:

$$\dot{\bar{g}}_N^+ = \varepsilon \dot{g}_N^- + \dot{g}_N^+ \tag{2.43}$$

$$\dot{\bar{g}}_N^+ \geq 0, \quad \Lambda_N \geq 0, \quad \dot{\bar{g}}_N^+ \Lambda_N = 0 \qquad \text{if } g_N = 0\,. \tag{2.44}$$

Proximal Point Formulation

The formulations of the contact and friction laws Eqs. (2.38), (2.39) and (2.40) are self-evident from a mechanical point of view, but are not suitable for numerical

computation. A more appropriate formulation can be attained using the proximal point to a convex set $C \in \mathbb{R}^{\mathbb{N}}$ [170], which is in general defined by:

$$\mathbf{prox}_C(\boldsymbol{x}) = \arg\min_{\boldsymbol{x}^* \in C} \|\boldsymbol{x} - \boldsymbol{x}^*\|, \quad \boldsymbol{x} \in \mathbb{R}^{\mathbb{N}} .$$

Then, the relations (2.38), (2.39) and (2.40) have the form [5, 6]

$$\lambda_B = \mathrm{prox}_{C_B}(\lambda_B - r\, g_N) , \tag{2.45}$$
$$\Lambda_B = \mathrm{prox}_{C_B}(\Lambda_B - r\, \dot{g}_B^+) , \tag{2.46}$$

$$\lambda_U = \mathrm{prox}_{C_U}(\lambda_U - r\, g_U) , \tag{2.47}$$
$$\Lambda_U = \mathrm{prox}_{C_U}(\Lambda_U - r\, \dot{g}_U^+) , \tag{2.48}$$

$$\boldsymbol{\lambda}_T = \mathbf{prox}_{C_T(\lambda_N)}(\boldsymbol{\lambda}_T - r\, \dot{\boldsymbol{g}}_T) , \tag{2.49}$$
$$\boldsymbol{\Lambda}_T = \mathbf{prox}_{C_T(\Lambda_N)}(\boldsymbol{\Lambda}_T - r\, \dot{\boldsymbol{g}}_T^+) . \tag{2.50}$$

with the corresponding convex sets being specified by:

$$\begin{aligned} C_B &= \{x \in \mathbb{R}\} , \\ C_U &= \{x \in \mathbb{R} |\ x \geq 0\} , \\ C_T(y) &= \{\boldsymbol{x} \in \mathbb{R}^2 |\ \|\boldsymbol{x}\| \leq \mu |y|\} \qquad \text{with } y \in \mathbb{R} . \end{aligned}$$

The variable $r > 0$ is an independent auxiliary parameter for each contact, which is arbitrary from the mechanical but not from the numerical point of view. The optimal choice of r with respect to numerical efficiency and stability of the solution scheme is discussed by FÖRG [62, 64].

2.3 Hydraulic Systems

Hydraulic systems are considered to be a network of basic components such as valves, hydraulic lines, and hydraulic cylinders. Abstracted to a logical structure, hydraulic systems comprise nodes and links. The behaviour of nodes is described by the pressure p as a possible state variable and volume V. Links that connect hydraulic volumes or mechanical systems are characterised by their volume flow Q as a possible state variable. Considering compressibility of the fluid, mass conservation results in a non-linear differential equation for nodes:

$$\dot{p} = \frac{E}{V} \left(\sum_j Q_j + Q_{\mathrm{mec}} \right) . \tag{2.51}$$

Therein Q_j denotes the hydraulic volume flow into the node and $Q_{\mathrm{mec}} = \boldsymbol{W}_{\mathrm{A,Q}}\, \boldsymbol{u}$ the variation of the volume due to mechanical systems such as hydraulic pistons. $\boldsymbol{W}_{\mathrm{A,Q}}$

represents the area of a mechanical interface to the hydraulic volume projected in the direction of the generalized velocity $\boldsymbol{u}$ of a connected multi-body system, yielding the generalised hydraulic force $\boldsymbol{h}_{\mathrm{hmi}}$. The bulk modulus $E = -V\frac{\partial p}{\partial V}$ depending on the air content and the pressure describes the compressibility of the fluid.

Hydraulic links with cross section area A and length l relate the change of flow $\dot{Q}$ with the connected pressures (p_{in}, p_{out} and pressure loss Δp) using momentum conservation:

$$\dot{Q} = \frac{1}{\rho l} A \left(p_{\mathrm{in}} - p_{\mathrm{out}} - \Delta p \right) \qquad (2.52)$$
$$\text{with} \quad \Delta p^k \propto Q \,.$$

The exponent k, which relates pressure loss and flow, is equal to 1 for laminar flow (Hagen-Poiseuille) and 0.5 for energy conservation according to Bernoulli's principle. Combining the mechanical and hydraulic equations, the differential equations for multi-body systems including first order systems are written as:

$$\begin{pmatrix} \boldsymbol{M} & 0 & 0 \\ 0 & \boldsymbol{M}_{\mathrm{Q}} & 0 \\ 0 & 0 & \boldsymbol{M}_{\mathrm{p}} \end{pmatrix} \begin{pmatrix} \dot{\boldsymbol{u}} \\ \dot{\boldsymbol{p}} \\ \dot{\boldsymbol{Q}} \end{pmatrix} = \begin{pmatrix} \boldsymbol{h} + \boldsymbol{h}_{\mathrm{hmi}} \\ \sum \boldsymbol{Q}_j + \boldsymbol{W}_{\mathrm{A,Q}}\, \boldsymbol{u} \\ \boldsymbol{p}_{\mathrm{in}} - \boldsymbol{p}_{\mathrm{out}} - \Delta \boldsymbol{p} \end{pmatrix} \qquad (2.53)$$

with the diagonal matrices $M_{\mathrm{Q},ii} = \frac{V_i}{E_i}$ and $M_{\mathrm{p},ii} = \frac{\rho_i l_i}{A_i}$. The structure of the equations of multi-body systems with and without hydraulic components is equal and therefore, the same integration schemes and solvers can be used.

One main problem of the simulation is that for small volumes and nearly incompressible fluids such as oil this approach leads to very stiff differential equations. Hence, the differential equation Eq. (2.51) can be substituted by an algebraic equation assuming an incompressible hydraulic volume. For an incompressible node only the mass conservation equation has to be satisfied as a constraint:

$$\sum_j Q_j + Q_{\mathrm{mec}} = 0 \,. \qquad (2.54)$$

3 Integration Schemes

Sophisticated computational methods have been established to adopt non-smooth mechanical models to a wide range of industrial applications. Additionally, the choice of the mechanical model, as well as the appropriate time-integration method, is essential. In order to integrate multi-body systems with rigid contacts, two different numerical methods can be distinguished: event-driven and time-stepping schemes.

3.1 Event-Driven Integration Schemes

Event-driven schemes [161] detect changes of the constraints—for example, closing of unilateral contacts or stick-slip transitions—and resolve the exact transition times. Between these events the motion of the system is smooth and can be computed by a standard ODE- or DAE-integrator, respectively. Hence, the performance of this type of integration scheme depends on the underlying integrator.
While the event-driven integration is very accurate, the detection of events can be time consuming, especially in the case of frequent transitions for example for systems with numerous contacts. This approach is best suited for systems with few impacts and changes of the system configuration.

3.2 Time-Stepping Integration Schemes

Time-stepping schemes are based on a time-discretization of the system dynamics including the contact conditions. Nowadays a wide variety of time-stepping methods for multi-body systems with unilateral constraints exist; see literature references [73, 74, 132, 185, 186]. Time-stepping schemes provide solutions for the average overall system dynamics and do not focus on single events. The number of combinatorial problems is reduced and event detection is avoided, resulting in a robust integration scheme for non-smooth systems with a large number of discontinuities. However, the common time-discretizations are of order one with fixed time-step size. In order to overcome these disadvantages, an integration scheme with step size selection and higher order integration based on time-stepping and extrapolation methods is developed within this work. This new approach is presented in the next chapter.
A half-explicit time-stepping algorithm of first order with constraints on the velocity level and with constant step size H is used as the base integration scheme. More details are provided by Förg [64]. A single integration step $l \to l+1$ is outlined below:

1. Compute the new generalised positions. $\boldsymbol{q}^{l+1} = \boldsymbol{q}^{l} + \boldsymbol{\Upsilon}^{l}\, \boldsymbol{u}^{l}\, H$.
2. Compute the distances $\boldsymbol{g}_U^{l+1} = \boldsymbol{g}_U(\boldsymbol{q}^{l+1}, t^{l+1})$ of all unilateral contacts.
3. Compute the index set $\{i : g_{U,i}^{l+1} \leq 0\}$ of active unilateral contacts.
4. Compute the generalised velocities by solving the discretized equations of motion considering the active constraints (index a) on velocity level:

$$\boldsymbol{u}^{l+1} = \boldsymbol{u}^{l} + \left(\boldsymbol{M}^{l+1}\right)^{-1} \left(\hat{\boldsymbol{h}}^{\,l+1} H + \boldsymbol{W}_a^{l+1}\, \boldsymbol{\Lambda}_a^{l+1}\right) , \tag{3.1a}$$

$$\dot{\boldsymbol{g}}_a^{l+1} = \dot{\boldsymbol{g}}_a(\boldsymbol{u}^{l+1}, \boldsymbol{q}^{l+1}, t^{l+1}) , \tag{3.1b}$$

$$\boldsymbol{\Lambda}_a^{l+1} = \mathbf{proj}(\dot{\boldsymbol{g}}_a^{l+1}, \boldsymbol{\Lambda}_a^{l+1}) . \tag{3.1c}$$

   Here, the half-explicit evaluation $\hat{\boldsymbol{h}}^{\,l+1} = \boldsymbol{h}(\boldsymbol{u}^{l}, \boldsymbol{q}^{l+1}, t^{l+1})$ is used to increase numerical stability. The function **proj** comprises the projection functions **prox** for bilateral and unilateral constraints as well as friction.
5. If constraints are violated, perform an additional drift correction step.

Details about this time integration scheme with a focus on efficient solution algorithms for the fix-point iteration of equation (3.1c) are provided by Förg [63,64].

Part II

Extrapolation Methods Based on Time-Stepping Schemes

4 Step Size Selection

Time-stepping schemes are well suited to mechanical systems including many set-valued force laws for bilateral and unilateral contacts with friction. For each time step the status of all contacts (open, closed, stick or slip) is determined and applied for the entire step. Assuming that all contacts remain unchanged during one integration step, the next state of motion is computed using the whole set of the discretized equations of motion and constraints. With a small step size, the discrete states and exact solution become similar. As a drawback, time-stepping methods, commonly used with constant time discretization, require a very small step size. In order to improve this, a strategy for selecting the step size within time stepping schemes is presented. A step size control with error estimation based on Richardson Extrapolation is used for the smooth part. Some heuristic adjustments are made since error estimation is not possible during the non-smooth part.

4.1 Step Size Control in ODE Theory

The aim of step size control in ODE theory is to adjust automatically the step size in order to achieve a prescribed tolerance of the local error. On the one hand the step size H_i must be chosen sufficiently small to yield the required precision of the computed results, and on the other hand the step size must be sufficiently large to avoid unnecessary computational work and round-off errors. The oldest device, used by Runge in his numerical examples, is to repeat the computations with half step size and compare the results. The underlying theory was presented by Richardson [169].

Consider a differential equation of the form

$$\dot{z} = f(t,z) \qquad \text{with initial values} \quad z(t_0) = z_0\,. \tag{4.1}$$

Suppose that with given initial values (t_0,z_0), step size H, and integration method of order p, a numerical approximation for $z(t_0 + H)$ is computed by one single integration step obtaining the result $\hat{z}^H(t_0+H)$. Using the same integration method and initial values, a second approximation $\hat{z}^{H/2}(t_0+H)$ by two successive integration (sub-)steps with step size $H/2$ is calculated. By comparing both numerical results, an estimation for the discretization error can be derived. Using asymptotic expansion, the discretization errors e^H and $e^{H/2}$ are written with the step size independent error

coefficient e_p:

$$\| \hat{z}^H - z \| = e^H = H^p e_p + \mathcal{O}(H^{p+1}) \tag{4.2a}$$

$$\| \hat{z}^{H/2} - z \| = e^{H/2} = \left(\frac{H}{2}\right)^p e_p + \mathcal{O}(H^{p+1}) \,. \tag{4.2b}$$

By combining Eq. (4.2a) and Eq. (4.2b), a first approximation for the discretization error is obtained:

$$\begin{aligned} e^{H/2}(t_0 + H) &= \| \hat{z}^{H/2}(t_0 + H) - z(t_0 + H) \| \\ &\approx \frac{1}{2^p - 1} \| \hat{z}^{H/2}(t_0 + H) - \hat{z}^H(t_0 + H) \| \,. \end{aligned} \tag{4.3}$$

With asymptotic expansion for $e_p(t_0 + H)$ and considering $e_p(t_0) = 0$ the error can be calculated as

$$e^{H/2}(t_0 + H) = \left(\frac{H}{2}\right)^p e_p(t_0 + H) + \mathcal{O}(H^{p+1}) \approx \frac{H^{p+1}}{2^p} e'_p(t_0) \,. \tag{4.4}$$

For an optimal step size H_{opt}, the error has to satisfy a prescribed tolerance Tol:

$$e^{H_{opt}/2}(t_0 + H_{opt}) = Tol = aTol + z \cdot rTol \,. \tag{4.5}$$

Solving Eqs. (4.4) and (4.5), an estimation for the optimal step size H_{opt} (Eq. (4.6)) can be calculated. The error coefficient e'_p can be estimated by Eqs. (4.3) and (4.4).

$$H_{opt} = \sqrt[p+1]{\frac{Tol}{e'_p} 2^p} = H \sqrt[p+1]{\frac{Tol}{\| \hat{z}^{H/2} - \hat{z}^H \|} (2^p - 1)} \tag{4.6}$$

For an efficient step size control, some common modifications are applied. Equation (4.6) is multiplied by a safety factor smaller than one so that the step will be accepted the next time with high probability. Furthermore, H_{opt} is not allowed to increase or decrease too fast.

4.2 Step Size Selection for Non-smooth Systems

Due to changing constraints and impacts coming along with discontinuous velocities, expansions in series are not valid and therefore the ODE theory is not applicable for non-smooth problems. Hence, two adjustments are made, namely adapting the error tolerances and "Gap Control".

4.2.1 Adaptation of Error Tolerances

As the requirements of ODE theory are violated by discontinuities in the velocity, an obvious first approach is to apply error estimation and step size control only to position parameters and to exclude the velocity parameters. A more sophisticated approach is to scale the velocity parameters. This approach is used in this work. For the purpose of error estimation and step size control, scaling variables with step size H according to their index is commonly used by several DAE solvers e.g. RADAU5 [87]. Thus, the velocity parameter u is multiplied by H and the position parameter q remains unchanged. Assuming $p = 1$, the inequality condition for an accepted step can be reformulated:

$$\begin{aligned} \| \hat{u}^{H/2} - \hat{u}^{H} \| \cdot H &\leq aTol + H \cdot \hat{u}^{H/2} \cdot rTol \\ \| \hat{q}^{H/2} - \hat{q}^{H} \| &\leq aTol + \hat{q}^{H/2} \cdot rTol\,. \end{aligned} \tag{4.7}$$

Using Eq. (4.7), only the position increment $u \cdot H$ is considered for error estimation, instead of the velocity parameter. For further information and justification see [84].

4.2.2 "Gap-Control" for Inelastic Impacts

Theory

The second adaptation made is a discretization of time with respect to events such as impacts, resulting in a kind of low-level event detection. This idea can be demonstrated by a simple example: a single mass falling on the ground with inelastic impact (see Figure 4.1).

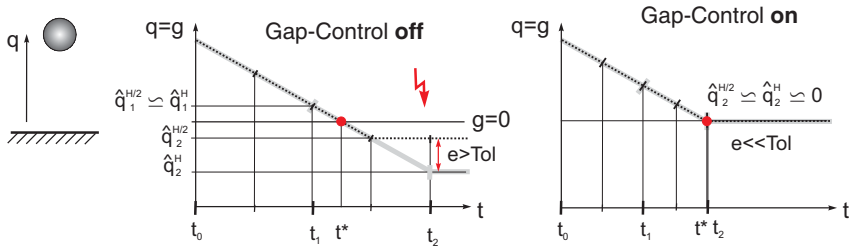


Figure 4.1: "Gap-Control": Falling mass with inelastic impact

Integration without "Gap-Control" is illustrated on the left side of Figure 4.1. Starting at t_0, two numerical approximations $\hat{q}_i^H$ (step size H) and $\hat{q}_i^{H/2}$ (two sub-steps with step size $H/2$) are calculated. Within the second integration step the impact occurs at $t = t^*$. As the time-stepping scheme identifies closing contacts by checking $g < 0$ at the beginning of each integration step, the contact is closed at $t = t_1 + H/2$ within the two sub-step approximation (dotted line) and at $t = t_2$ within the one

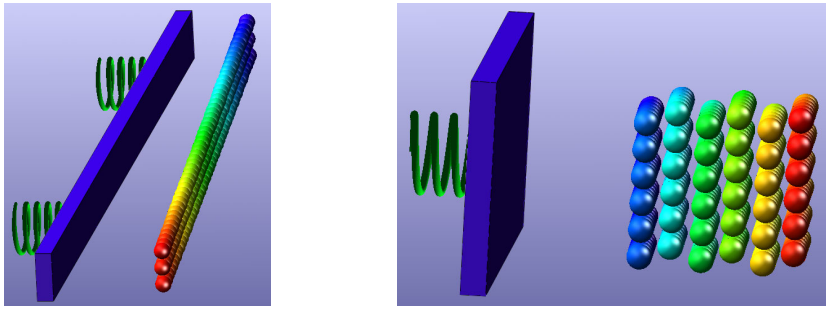
step approximation (full line). Besides the constraint violation, the error estimation by comparing $\hat{q}_i^H$ and $\hat{q}_i^{H/2}$ makes this integration step likely to be rejected.
As seen on the right in Figure 4.1, error and penetration can be reduced by changing the time discretization towards a kind of low-level event detection. The next impact t^* is estimated using the gap velocities $\dot{g}$. This information is then used to reduce the step size H_{i+1} for the next step, so that the impact is located at the end of the integration step. Hence, constraint violation and errors in the position parameters are reduced without resolving the events exactly.

The possibility exists that more than one contact will close in subsequent integration steps while integrating systems with several contacts. Therefore, a strategy on how to deal with multiple impacts at t_i^* is proposed. In this work three different strategies are used. Let $\{t_i^*\}$ denote all points in time at which an impact is estimated.

- **Strategy 1:** fastest integration progress by choosing the biggest t_i^*
- **Strategy 2:** choose the biggest t_i^* on the condition that penetration for all $t_j^* < t_i^*$ is smaller than a given tolerance
- **Strategy 3:** highest accuracy by choosing the smallest t_i^*

Discussion and Examples

To demonstrate robustness and efficiency of the presented method, even for systems with many constraints, two demonstrative problems with inelastic impacts are solved (see Figure 4.2). More application-related and complex numerical examples can be found at the end of this section.



(a) Test System 1: consecutive impacts

(b) Test System 2: simultaneous impacts

Figure 4.2: Test systems with multiple inelastic impacts

The system comprises 216 balls and a wall. Frictional contacts are possible between each ball and between ball and wall. The wall is inertially linked with a spring. The balls are shot against the wall with random velocities between 150 and 200m/s leading to multiple impacts. The system can be described by 650 degrees of freedom and 23436 possible contacts. Tolerances $aTol = 10^{-7}$ and $rTol = 10^{-5}$ are used.
Figure 4.3 illustrates the functionality of the "Gap-Control" algorithm. The impacts (red dots) and time discretization (vertical lines) with and without "Gap-Control"

are plotted for test system 1. Without "Gap-Control" the points in time for the impacts are scattered almost arbitrarily, whereas with "Gap-Control" nearly every impact occurs right before the end of a single integration step.

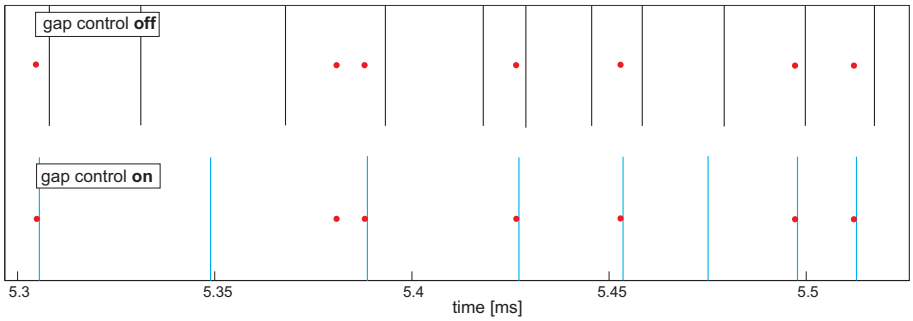


Figure 4.3: Time discretization adjusted by "Gap-Control"

Table 4.1: Test system 1: Effect of adapted error test and "Gap-Control"

scale $\boldsymbol{u}$ for error test	"Gap - Control"	integration steps (comput. effort)	refused steps	mean penetration
no	off	1000	197	3.1e-6
no	on	1065	159	1.2e-8
yes	off	446	45	9.3e-6
yes	on	424	3	9.6e-8
const. step size	H= 1e-7	4106	-	3.2e-6
const. step size	H= 3e-7	1370	-	9.0e-6

The effects of adapting the error test and "Gap-Control" for both test systems can be studied in Tables 4.1 and 4.2. The individual columns represent the number of integration steps and refused steps, as well as the geometric mean value of constraint violations due to penetration. For better legibility, the number of integration steps is normalized. The results from time-stepping with constant step size can be found in the last two rows of Tables 4.1 and 4.2. To make results comparable, the number of integration steps with constant step size is divided by three, as there are three sub-steps necessary with step size selection compared to one without. Hence, the number of integration steps represents the computational effort.
Taking penetration as a measure for accuracy we can compare the results with and without step size selection. The number of integration steps is reduced significantly. For both examples "Gap-Control" has two positive effects: the number of refused steps and penetration are both reduced, whereas the total number of steps remains nearly unchanged. The data is obtained with "Gap-Control" strategy 1. Penetration can be reduced to approximately 10^{-11} by strategy 3.

Table 4.2: Test system 2: Effect of adapted error test and "Gap-Control"

scale $\boldsymbol{u}$ for error test	"Gap - Control"	integration steps (comput. effort)	refused steps	mean penetration
no	off	1000	541	6.7e-8
no	on	1032	268	4.6e-9
yes	off	257	116	7.4e-7
yes	on	360	57	7.6e-8
const. step size	H= 1e-6	2808	-	5.6e-8
const. step size	H= 2e-7	14045	-	4.3e-9

4.2.3 "Gap-Control" for Elastic Impacts

For elastic impacts, the post-impact velocity $\dot{g}_N^+ = -\varepsilon \dot{g}_N^-$ is in general not equal to zero. Hence, the "Gap-Control" strategy to place the impact at the end of the integration step does not work. The idea of "Gap-Control" in the elastic case is to use very small step sizes for integration steps with impacts. Again, $\dot{g}$ is used to estimate the next impact t^*. The step size is reduced to H_1^{GC} if t^* is within the next integration step with step size H^{ssc} (cp. Figure 4.4). H^{ssc} is the step size suggested by the step size control. The integration is stopped before the impact occurs and H_1^{ssc} is stored. Next, the integration step containing t^* is performed with H_2^{GC}. After the impact, integration proceed with the stored step size H_1^{ssc}. The value for H_2^{GC} relative to the previous step size is estimated using the applied integration tolerances as a measure for the accuracy of the impact estimation t^*.

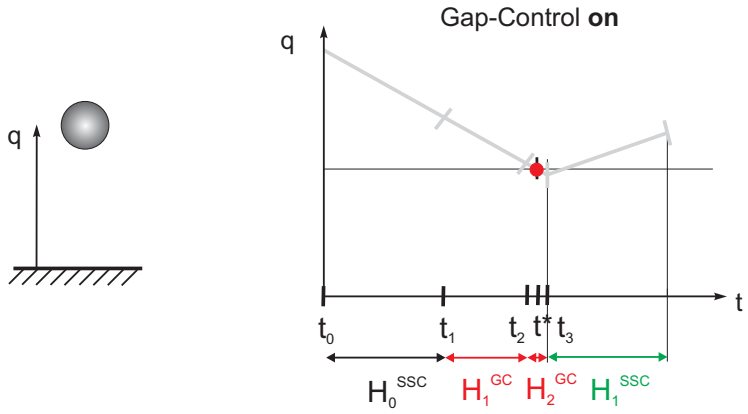


Figure 4.4: "Gap-Control": Falling mass with elastic impact

Using this strategy for the example of the falling mass, the number of integration steps could be reduced from 807 to 638 steps and the number of refused steps from

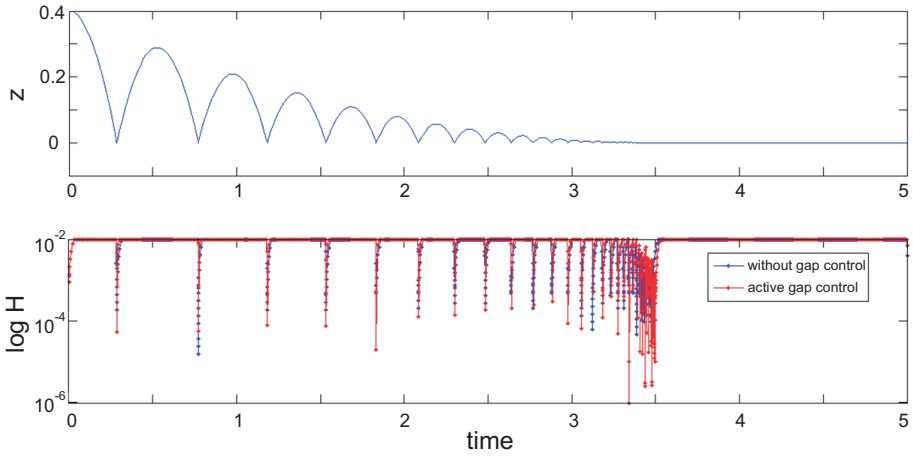


Figure 4.5: Falling mass with elastic impacts: position and step size

208 to 48 (cp. Figure 4.5). The results for the test systems of the previous section are shown in Table 4.3. Again, the effect of "Gap-Control" is more towards accuracy in terms of penetration than towards reduction of the integration steps.

Table 4.3: Test system 1 and 2: Effects of "Gap-Control"

"Gap - Control"	integration steps	refused steps	mean penetration
Test system 1			
off	1000	257	5.9e-5
on	790	39	5.3e-7
Test system 2			
off	819	386	4.2e-6
on	1000	134	2.2e-7

Although the adaptations presented in this section provide a robust and efficient integration scheme, it is important to point out that for integration steps with discontinuities such as closing contacts or stick slip transitions, an error estimation in terms of ODE theory is not possible. It cannot be assured that the integration error is lower than a prescribed tolerance. For the non-smooth part, step size selection is primarily directed towards robustness and stability rather than accuracy.
For higher accuracy requirements "Gap-Control" can be used to resolve nearly every impact, achieving a low-level event detection.

5 Extrapolation Schemes for Non-smooth Systems

5.1 Extrapolation Methods for ODEs

In this section a brief introduction to extrapolation methods is given. More detailed information can be found in [39, 40, 85, 187]. In this work extrapolation is used to increase the integration order in the smooth part. The use of time-stepping as a base integration scheme for extrapolation can also be found in STUDER et al. [188].

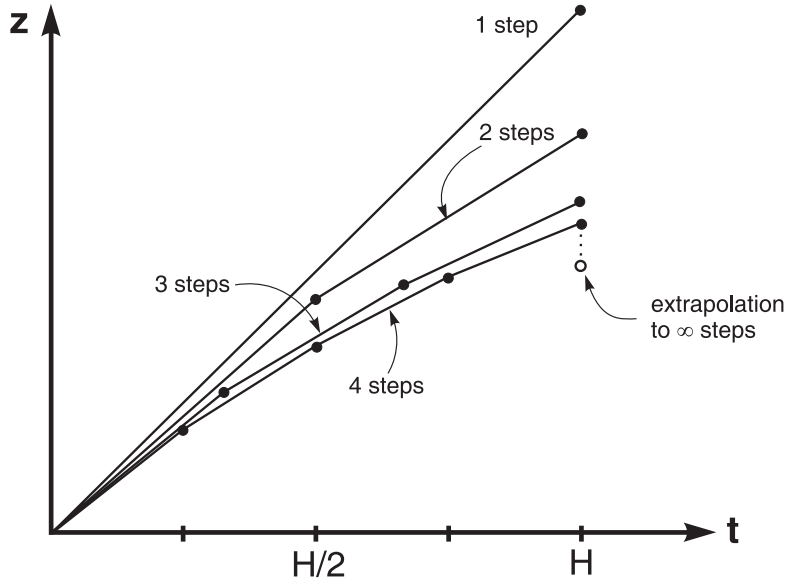


Figure 5.1: Extrapolation methods

The idea described in RICHARDSON [169] is to consider the final answer of a numerical calculation as itself being an analytic function with step size H as parameter. That function can be probed by performing numerical calculations with various step sizes, none of them necessarily small enough to yield the desired accuracy. Different sequences of fine and finer sub-steps span the large interval H. The results are fitted to an analytical function and extrapolated to $H \to 0$ giving an answer that is supposed to correspond to infinitely fine sub-steps (see Figure 5.1).
Considering a base integration scheme of order p and using k approximations calculated by the base integration scheme with different sub-step sizes, extrapolation methods allow the construction of an order $p + k - 1$ approximation $\tilde{z}(t_0 + H)$.

First we choose k different sequences of sub-steps and corresponding step sizes h_i:

$$h_i = \frac{H}{n_i} \qquad \text{with} \quad i = 1...k \quad \text{and} \quad n_1 = 1 < n_2 < ... < n_k \quad (n_i \in \mathbf{Z}) \,.$$

Then, the numerical results of the initial value problem Eq. (4.1) are computed for all sequences to obtain k estimates:

$$\hat{z}^{h_i}(t_0 + H) \qquad i = 1...k \,. \tag{5.1}$$

Calculating the interpolation polynomial

$$P(H) = \tilde{z} - e_p H^p - e_{p+1} H^{p+1} - ... - e_{p+k-2} H^{p+k-2} \tag{5.2}$$

derived by an asymptotic expansion of the global error and using the estimates $\hat{z}^{h_i}(t_0 + H) = P(h_i)$ for all sequences, the error coefficients e_p in Eq. (5.2) are eliminated. Finally, "extrapolation to the limit" $h_i \to 0$ is done to get $P(0) = \tilde{z}$ as the extrapolation result of order $p + k - 1$.

Example
This example illustrates how to build the extrapolation tableau for order $p = 1$ and two sub-step series ($k = 2$). Two approximations $\hat{z}^H(t_0 + H)$ and $\hat{z}^{H/2}(t_0 + H)$ are calculated using the (sub-)step sizes $h_1 = H$ and $h_2 = H/2$:

$$\begin{array}{lcccl} z(t_0) & \xrightarrow{\qquad\qquad H \qquad\qquad} & & & \hat{z}^H(t_0 + H) \\ z(t_0) & \xrightarrow{H/2} & \hat{z}^{H/2}(t_0 + H/2) & \xrightarrow{H/2} & \hat{z}^{H/2}(t_0 + H) \end{array}$$

$$\begin{array}{lccl} \text{interpolation polynom :} & P(H) & = \tilde{z} - e_p H \\ \text{supporting points:} & \hat{z}^H & = \tilde{z} - e_p H \\ & \hat{z}^{H/2} & = \tilde{z} - e_p \frac{H}{2} \\ \\ \text{solving for } \tilde{z} : & \Rightarrow & \tilde{z} = 2 \cdot \hat{z}^{H/2} - \hat{z}^H \,. \end{array}$$

Higher order extrapolations can be derived analogically:

$$\begin{aligned} \text{order 3:} \quad \tilde{z} = & \frac{1}{3} \cdot \hat{z}^H - 2 \cdot \hat{z}^{H/2} + \frac{8}{3} \cdot \hat{z}^{H/4} \\ \text{order 4:} \quad \tilde{z} = & -\frac{1}{15} \cdot \hat{z}^H + \hat{z}^{H/2} - \frac{16}{3} \cdot \hat{z}^{H/4} + \frac{27}{5} \cdot \hat{z}^{H/6} \\ \text{order 5:} \quad \tilde{z} = & \frac{1}{30} \cdot \hat{z}^H - 2 \cdot \hat{z}^{H/2} + \frac{27}{2} \cdot \hat{z}^{H/3} - \frac{64}{3} \cdot \hat{z}^{H/4} + \frac{54}{5} \cdot \hat{z}^{H/6} \,. \end{aligned}$$

5.2 Extrapolation Methods Applied to Time-Stepping Schemes

5.2.1 Inelastic Impacts

Time evolution of non-smooth systems can be divided into smooth parts and integration steps with changes in the structure of the underlying equation of motion, e.g. due to closing contacts or stick-slip transitions. It is important to point out that it is not feasible to apply any higher order integration scheme to non-smooth integration steps, because the asymptotic expansion of the integration error is only valid for continuous functions. Therefore, time integration during the non-smooth part is always performed with integration schemes with maximum order one. Extrapolation methods are applied during the smooth part that allows an increase in the integration order. For contacts with friction and plastic impacts, no adaptations have to be made. The extension to elastic impacts is presented in the next section.

Step Size Selection and Extrapolation

Accuracy: Variable Order or Variable Step Size
To achieve the desired accuracy of the numerical solution, extrapolation methods offer two possibilities, namely a variable step size and a variable integration order. For a variable step size, the integration step is recalculated using a reduced step size in the case the estimated error exceeds the specified tolerance. Within the concept of a variable order the integration order is increased by adding a new series of sub-steps to the extrapolation scheme until the estimated error is smaller than the specified tolerance.

Within this work a variable step size is used, coming along with a fixed order during the smooth part and order $p = 1$ during non-smooth integration steps.
From a computational and algorithmic point of view, an algorithm based on extrapolation methods with a variable step size, which itself contains algorithmic parallelism, has the advantage that they are almost ideal for the application of parallel computing techniques. Using a fixed order, all sub-step series can be calculated simultaneously by parallel computer architectures e.g. multi-core CPUs. In contrast, for variable order the sub-step series are consecutively evaluated until the desired accuracy is reached.
Facing the choice between variable order and variable step size from a numerical and mathematical point of view, variable step size is preferable. Extrapolation techniques are known to be sensitive against discontinuities. For high approximation orders the right side of the differential equation must be several times continuously differentiable. Modelling technical systems, one has to deal with higher order discontinuities even in the smooth part. Various reasons for these discontinuities, which are not necessarily obvious at first glance, include force inputs from discrete measurements, function evaluations by table look-up, discrete controllers, and numerical

evaluations of contact kinematics between non-analytical contours. For these reasons increasing the integration order is restricted in practice. It cannot be assured that extrapolation is always contractive. Therefore, for a robust algorithm, the best approach is to choose a fixed integration order appropriate for the modelled system and reduce the step size to achieve the desired accuracy instead of increasing the integration order.

Error Estimation: Extrapolation or Embedded Methods

To estimate the integration error two approximations obtained with different step sizes are compared as described in previous sections. For the smooth part, there are two different possibilities to construct a step size control: by extrapolation (comparing two approximations obtained by different sub-step size (Figure 5.2)) or embedded methods (comparing two approximations obtained by different integration orders (Figure 5.3)). Both possibilities are illustrated in Figures 5.2 and 5.3 for order 2.

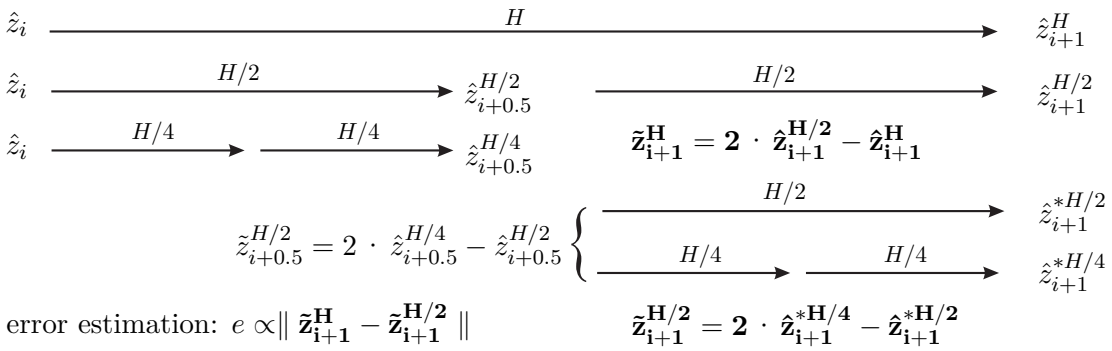


Figure 5.2: Integration order 2 and step size control by extrapolation

As previously mentioned, extrapolation to higher integration orders is sometimes not reliable in practice. Therefore, step size control by comparing two approximations obtained by different step sizes seems to be preferable. This assumption is confirmed by numerical examples in the next section. As shown in Figure 5.3, the embedded technique requires one sub-step less. Using parallel computing techniques, this reputed drawback of extrapolation methods can be eliminated.

A flow chart of one integration step is depicted in Figure 5.4. Starting with t_i, $\hat{z}_i$ and H as the step size for the overall integration step, the first half of the sub-steps is computed. If there are any events detected within the interval from t_i to $t_i + H/2$, the integration order is restricted to one. By comparing the result of the smallest sub-step with the next smallest sub-step, a decision must be made whether the half-step is accepted or refused. In the case of unchanged constraints, all remaining (sub-)steps are computed. Again, it has to be checked if there are any events within the integration step. Only if there are no events, the extrapolation scheme for higher order approximation can be applied.

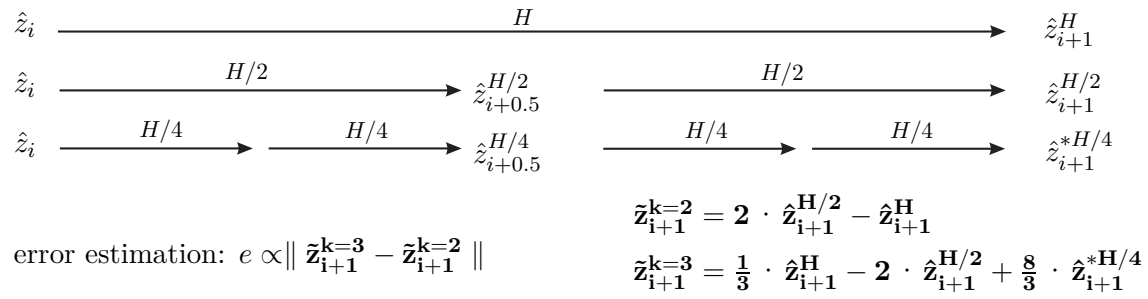


Figure 5.3: Integration order 2/3 and embedded step size control

5.2.2 Failure of Extrapolation: Elastic Impacts

According to STUDER et al. [188] problems with extrapolation can occur in cases of enduring unilateral contacts which are described by Newton's impact law:

$$\dot{g}_N^+ = -\varepsilon \dot{g}_N^- \,. \tag{5.3}$$

This equation must hold for both impact and enduring contacts. Consider a ball falling onto a table with restitution coefficient $0 < \varepsilon < 1$. If the ball touches the table, the velocity of the ball is inverted according to the impact law Eq. (5.3). After a while the ball will remain on the table yielding an enduring contact. Applying now the impact law to one integration step with step size H starting with $\gamma^- = \dot{g}_N^-$, yields $\gamma^+ = -\varepsilon\gamma^-$. If the same integration progress is done with a refined time grid of $H/2$, the result is $\gamma^+ = \varepsilon^2\gamma^-$, which is a completely different result and even a different sign. Eq. (5.3) is independent of the step size. Hence, the impact law must be seen as a mapping, and not as a consistent integration scheme for enduring contacts. As $\varepsilon < 1$, this mapping is a contraction and therefore noise amplification is normally not observed.

When using the time-stepping integration scheme as base integrator for extrapolation methods, the mapping given by Newton's impact law is no longer always a contraction. For example, extrapolation to order 2 with step size H and $H/2$ is done:

$$\begin{array}{lcccl} \gamma^- & \xrightarrow{\qquad\qquad H \qquad\qquad} & & & -\varepsilon\gamma^- \\ \gamma^- & \xrightarrow{H/2} & -\varepsilon\gamma^- & \xrightarrow{H/2} & \varepsilon^2\gamma^- \end{array}$$

$$\text{Extrapolation: } \tilde{\gamma}^+ = \underbrace{(2\varepsilon^2 + \varepsilon)}_{>1 \forall \varepsilon > 0.5} \gamma^- \tag{5.4}$$

For $\varepsilon > 0.5$ the effective restitution coefficient $2\varepsilon^2 + \varepsilon$ is greater than one. This may lead to noise amplification, growing kinetic energy, and gap velocities resulting in

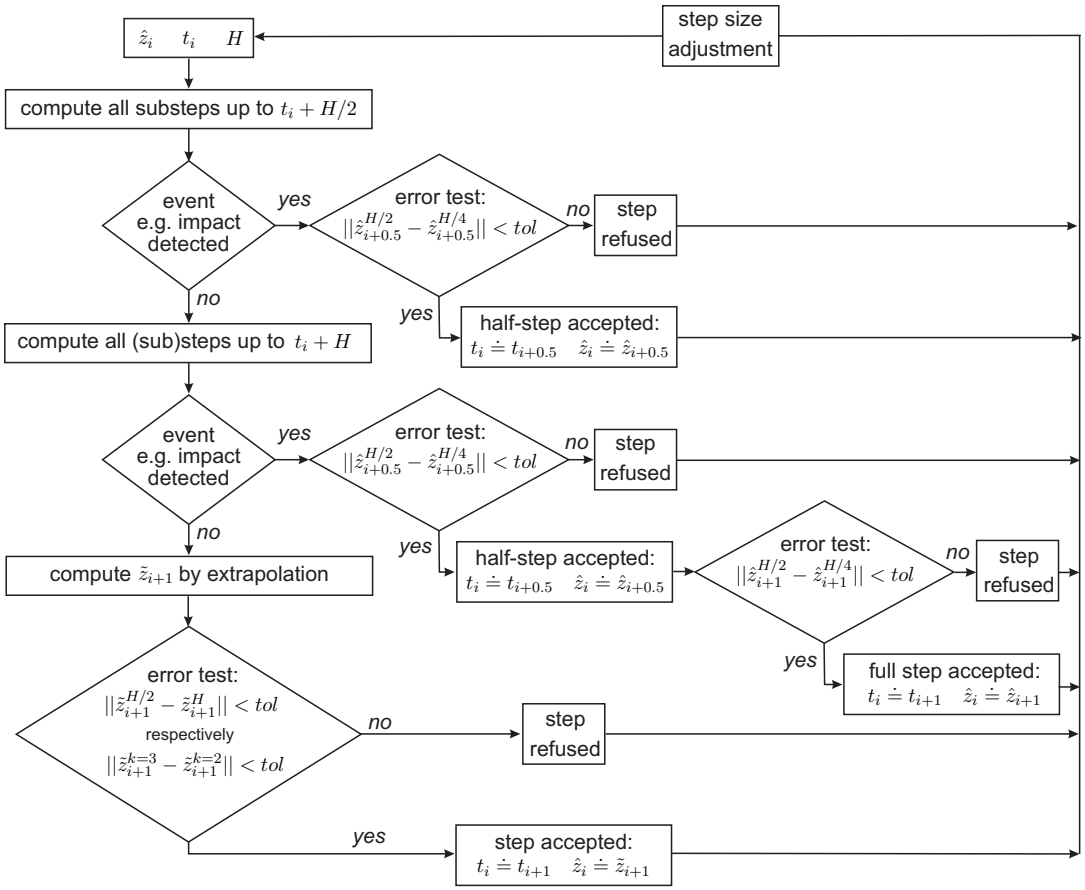


Figure 5.4: Flow chart of one integration step

contacts that may open after a while. In Figure 5.5 the amplification for different extrapolation schemes is plotted as a function of the restitution coefficient.

Modifications to Prevent Failure of Extrapolation

There are two approaches to prevent extrapolation from failure. Studer [188] demonstrated that if extrapolation is restricted to an uneven number of sub-steps, then the extrapolated mapping remains contractive.
The second approach is to set the restitution coefficient to zero for enduring contacts. This approach is used in this work. It is done by implementing a velocity dependent restitution coefficient that is smooth set to zero for gap velocities becoming smaller than a predefined limit. The effect can be seen in Figure 5.6 where the simulation results for a bouncing ball are shown. For extrapolation without modification (left side), the contact opens due to the described numerical reasons. With modifications, the contact remains closed and no oscillations can be observed.

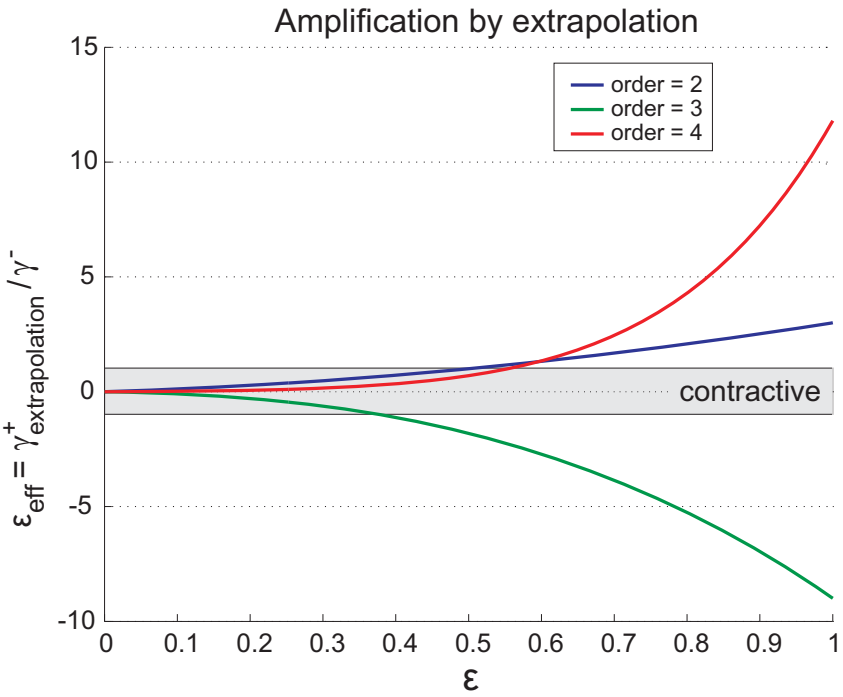


Figure 5.5: Effective restitution coefficient

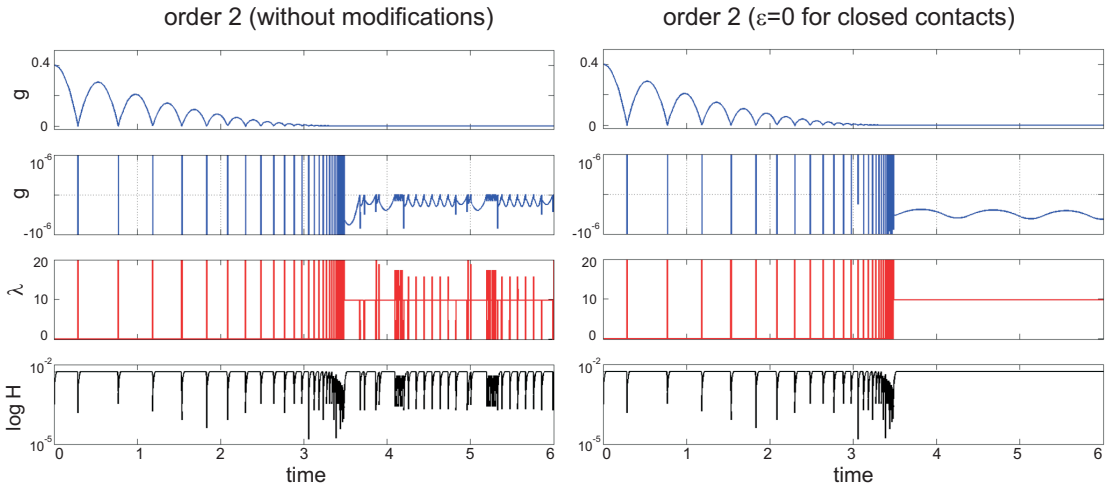


Figure 5.6: Bouncing ball

5.3 Examples

All examples are computed using the software MBSim [206] that was developed at the Institute of Applied Mechanics [1], Technische Universität München. The software is freely available under the GNU Lesser General Public License.

5.3.1 Woodpecker Toy on Elastic Pole

The woodpecker toy is a classic example for non-smooth multi-body system dynamics (see Figure 5.7). Hammering down a pole, the system uses gravity as its energy source and stick-slip effects between the guiding sleeve and the pole as the switch between free and constrained motion. With the sleeve sticking, the spring forces the woodpecker back towards the pole. The pole is modelled by a spatial bending beam using finite-element formulation for slender structures and is based on Euler-Bernoulli's theory. The tolerance between pole and sleeve is 0.8mm. Overall the system has 29 degrees of freedom, six for the plane motion of the woodpecker and the sleeve and 23 for the elastic beam as well as six frictional contacts, one for the beak and four between the edges of the sleeve and the pole. For a more detailed description of the model see [63].

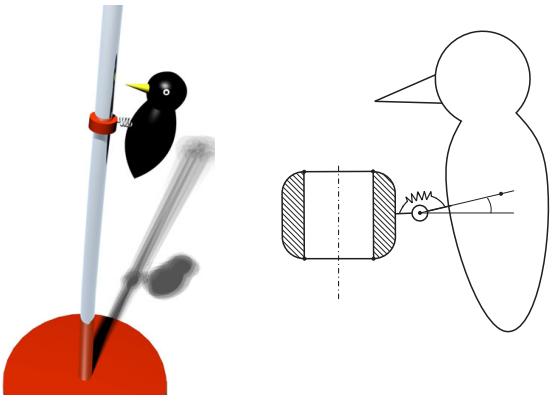


Figure 5.7: Model of a woodpecker toy

Figure 5.8 shows the data obtained with a maximal integration order two, step size selection by extrapolation and "Gap-Control" strategy 1. The nearly horizontal lines in the diagram for the vertical sleeve position indicate sticking between the sleeve and the pole. A closer look shows that although the contacts remain closed during such periods, there are stick-slip transitions due to vibrations of the elastic pole indicated by integration order one.

Table 5.1 compares the different methods of error estimation for varying orders. As the possible contact point between beak or sleeve and pole, mathematically described by shape functions, is solved numerically with a certain tolerance, there are

1 http://www.amm.mw.tum.de

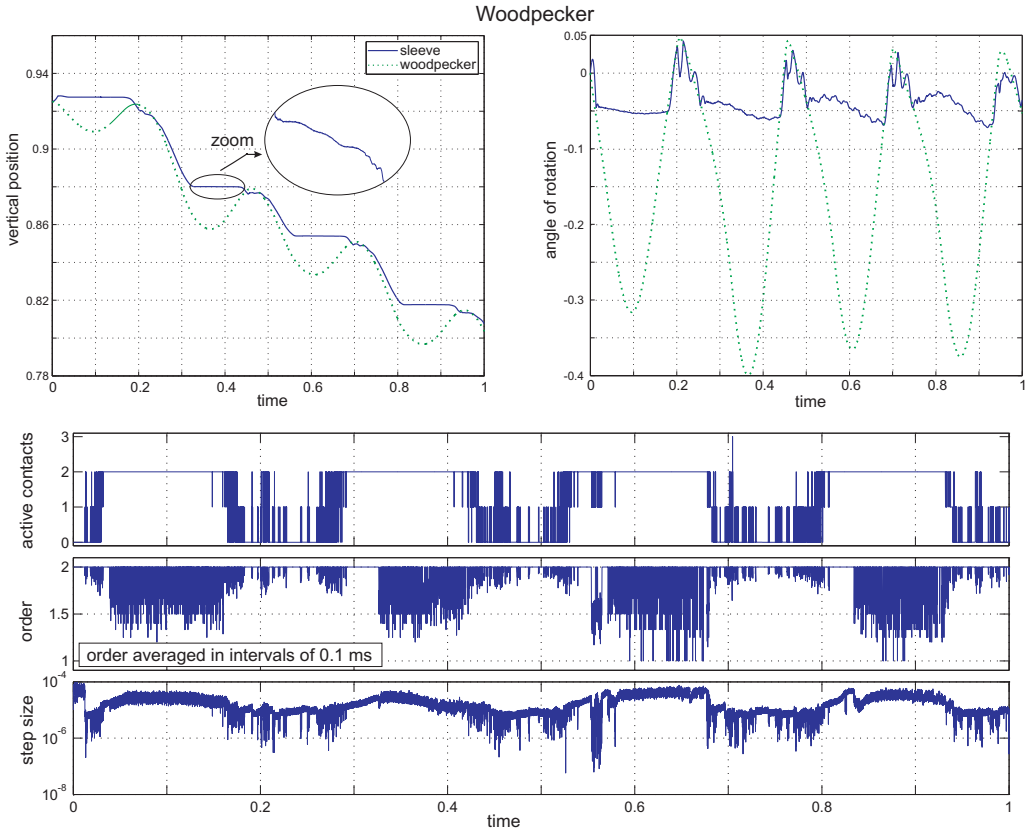


Figure 5.8: Woodpecker Toy: Results

Table 5.1: Woodpecker Toy: Results

Step Size Selection by Extrapolation

order	computational time (QuadCore)	integration steps	steps refused	function evaluations
1	184s	1000	21	3062
2	154s	274	51	2481
3	246s	222	59	4419

Embedded Step Size Selection

order	computational time (QuadCore)	integration steps	steps refused	function evaluations
1	256s	1381	39	4261
2	156s	332	25	2414
3	306s	439	98	6232
4	277s	267	96	5203

weak discontinuities, even in the smooth part. According to the results (Table 5.1), extrapolations to higher orders than two cannot increase computational efficiency. Comparing the results with respect to the presented methods of error estimation, step size selection by extrapolation seems to be more efficient than embedded step size selection. To improve legibility the number of integration steps is normalised.

5.3.2 Lottery

The German Lotto 6/49 has a random drawing of six numbers selected from numbers 1 to 49, using 49 balls mixed in a barrel (see Figure 5.9). The player picks six numbers between 1 and 49 and wins a prize if three or more numbers match the numbers on the balls drawn from the barrel.
Using six degrees of freedom for each ball, the equation of motion has 294 degrees of freedom. Between all bodies spatial Coulomb friction with a friction coefficient of $\mu = 0.2$ is considered. Therefore, up to 2156 possible unilateral constraints must be handled. At the beginning of the simulation all balls are outside the barrel on a plate with inclination. Due to gravity the balls roll and bounce into the barrel. After approximately 1.5 seconds all balls are inside the turning barrel. Deflectors stuck to the barrel strengthen mixing.
Integration is carried out with step size selection by extrapolation, extrapolation to order two in the smooth part and without "Gap-Control".

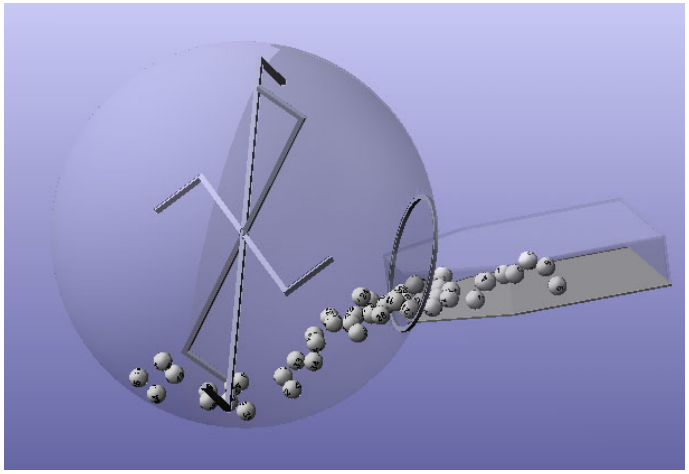


Figure 5.9: Lottery

The results are shown in Figure 5.10. Most of the balls are in free flight at the beginning and only a few contacts are closed at the same time. Therefore, the majority of integration steps are performed with integration order two. After a while the balls are concentrated at the bottom of the barrel and nearly every ball is subject to a closed contact. Hence, nearly all integration steps are restricted to order one due to changing constraints. Due to a high number of changing constraints, the computational time could not be reduced by extrapolation to a higher integration

order for this example. The uniform distribution of the step size demonstrates the robustness of the derived integration scheme.

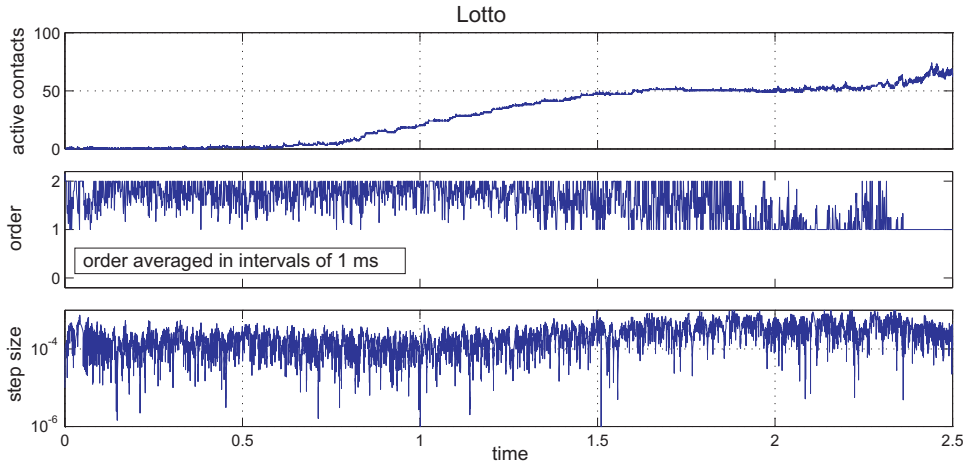


Figure 5.10: Lottery: Results

5.4 Parallel Computing

Frequency scaling was the dominant reason for improvements in computer performance from the mid-1980s until 2004. Increasing processor power consumption led to the end of frequency scaling as the dominant computer architecture paradigm in 2004 and multiprocessor architectures for parallel computing became popular [61]. Therefore, parallel computing is one of the most promising possibilities to accelerate numerical integration and speed up multi-body simulations.
As mentioned in the previous section, extrapolation methods themselves contain algorithmic parallelism and therefore are ideal for the application of parallel computing techniques. All sub-step series can be calculated simultaneously by parallel computer architectures e.g. multicore CPUs. In this work *OpenMP* (Open Multi-Processing) [1, 33] is used for parallelising the sub-step series. *OpenMP* is an application programming interface for shared memory multiprocessing programming. It consists of a set of compiler directives, library routines, and environment variables. Parallel threads are specified through the use of compiler directives and therefore can be embedded with little effort in existing source codes.

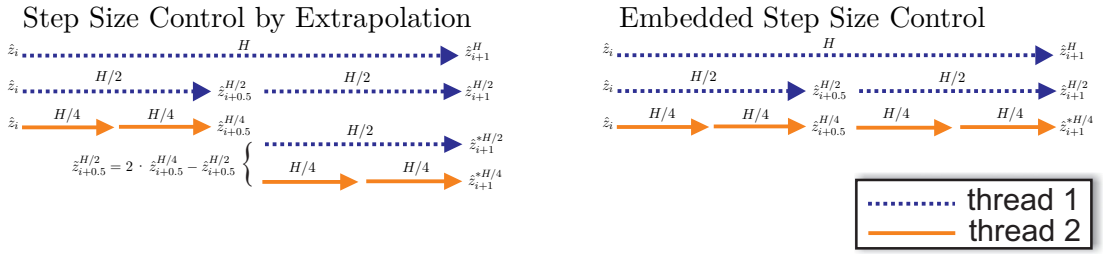


Figure 5.11: Parallelisation: Sub-steps grouped by threads

For order two, Figure 5.11 exemplary illustrates how single sub-steps are grouped by threads for parallelisation. On the left side, the sub-steps for integration with step size control by extrapolation are depicted. Eight sub-steps can be grouped into two parallel threads, each with four subsequent steps yielding an acceleration of $8 : 4 = 2$ on a dual-core CPU. On the right side of Figure 5.11 (embedded step size control) seven subsequent steps are grouped into two parallel threads with four and three subsequent steps, respectively, yielding an acceleration of $7 : 4 = 1.75$. In Table 5.2 the theoretical accelerations for different integration orders are calculated.

The derived time-stepping scheme with step size selection and extrapolation to higher integration order has been extensively tested using various numerical examples with academic (e.g. the woodpecker toy [206]) and industrial backgrounds (e.g. the AVS valve train [108] or a pushbelt CVT [176], Figure 5.12). The achieved accelerations due to parallelisation are shown in Table 5.4 and are very close to the theoretical maximum values.

Table 5.2: Acceleration due to parallelisation (theory)

Step Size Control by Extrapolation

maximal order	number of threads	theoretical acceleration
1	2	1.5
2	2	2
3	3	2.83

Embedded Step Size Control

maximal order	number of threads	theoretical acceleration
1	2	1.5
2	2	1.75
3	3	2.16
4	3	2.66

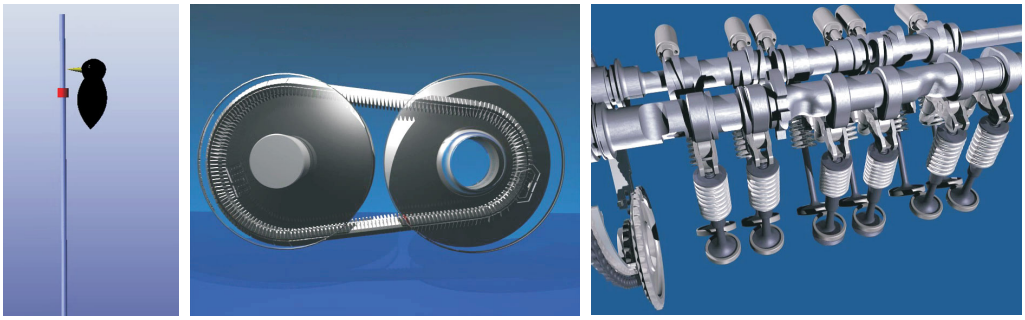


Figure 5.12: Examples for testing the derived integration scheme

Table 5.3: Acceleration due to parallelisation (numerical studies)

maximal order	number of threads	acceleration (theory)	acceleration (practice)
1	2	1.5	1.4 - 1.5
2	2	2	1.8 - 1.9
3	3	2.83	2.4 - 2.6

Part III

Dynamics of Variable Valve Trains

Dynamics of Variable Valve Trains: Outline

Variable valve trains provide a comprehensive means of adjustment in terms of variable valve timing and valve lift. Efficiency, torque, and emissions of gasoline engines are improved, but in return an increasingly complex mechatronic system must be handled. For optimization and design of such kinds of complex systems, detailed simulation models covering different physical domains such as mechanics, hydraulics, electrodynamics, and control are indispensable.
Topic of this work is the modelling and simulation of the variable valve train Audi valvelift system (AVS) of the Audi 2.8 litre V6 FSI engine. To investigate the dynamical behaviour and interactions between all subsystems e.g. between chain vibrations and shifting events, an overall simulation model of the timing assembly comprising chain drive, chain tensioner, hydraulic cam phasing system, valve train, and AVS system including the electromagnetic actuator, is derived (cp. Figure 5.13).

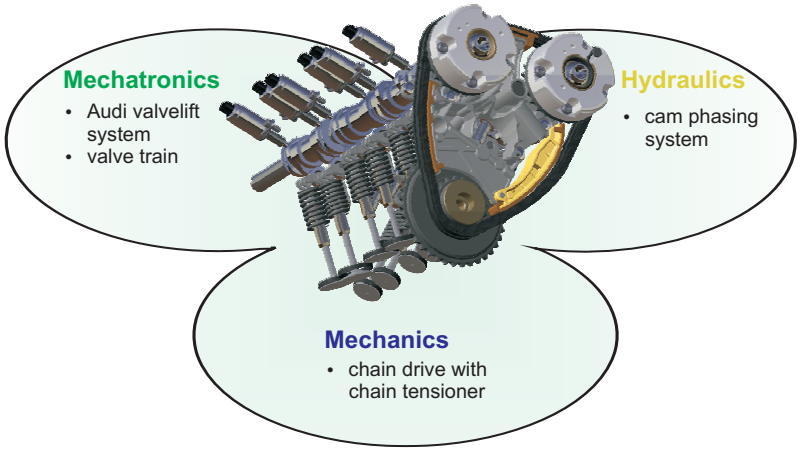


Figure 5.13: Timing assembly as multidisciplinary system

The first chapter (Chapter 6) of this section gives a short overview of variable valve timing and the design and concept of the Audi valvelift system. Moreover, modelling approaches for valve trains and the contact between cam and roller are presented.
In the next chapters, the components of the AVS valve train are described in detail and analysed by simulation and experiment. Specifically designed test rigs are used to validate the simulation models. The AVS specific switching unit is described in Chapter 7. Besides dynamical investigations, special friction measurements have been carried out. A simulation model of the electromagnetic actuator based on an analytical solution of Maxwell's field theory equations is derived in Chapter 8. The equations of motion for a mass carrying coil spring based on a curved beam, including coil to coil contacts, are presented in Chapter 9. Modelling of the hydraulic lash adjuster is shown in Chapter 10.

The dynamic of the entire timing assembly is analysed in Chapter 11. Models of the cam phasing system, the chain drive, and the chain tensioner have been adapted from previous works. Parallel co-simulation techniques [67] are used to analyse interactions between all components during a switching process.

6 Variable Valve Timing

New emissions legislation, tougher CO_2 standards, and increasing consumer expectations concerning fuel consumption make new solutions in engine technology necessary. Whereas in the past increased engine power, torque, and response characteristics have been the major development objectives, nowadays fuel efficiency and emissions are the most important factors in engine technology.
For gasoline engines, promising approaches for reducing emissions and increasing fuel efficiency include downsizing concepts, gasoline direct injection combined with stratified fuel charge, variable valve trains, engines with variable compression ratio, as well as the reduction of friction and weight. Downsizing is a well known and widely used technology to achieve higher fuel efficiency. With turbo- or supercharging concepts the engine displacement or the number of cylinders can be reduced resulting in a higher mean load and lower heat and frictional losses. The potential of downsizing is limited as the compression ratio has to be reduced to avoid engine knocking under full load. Hence, a variable compression ratio offers an opportunity to increase the possible downsizing potential. A higher compression ratio results in better fuel efficiency, but also has negative effects on losses due to friction. Flierl et al. [60] investigated the effects of variable compression with a single cylinder engine. Under partial load, fuel efficiency could be improved about 2.5% with a compression ratio of $\epsilon = 13$. Raising ϵ above 14 has no further positive effects on fuel consumption due to the higher friction level.
Along with direct injection, the variable valve train offers the largest potential for fuel consumption improvement gained via one individual measure. Compared to direct injection methods, variable valve train techniques have the advantage that with homogeneous stoichiometric combustion processes, conventional exhaust aftertreatment technologies can be used without requiring special fuel qualities. Besides fuel consumption, variable valve timing is increasingly used in production engines to improve torque characteristics, especially at low speed, engine response, and emissions of gasoline engines.

6.1 Valve Train Variability

In a reciprocating internal combustion engine, the valve train is responsible for controlling the gas exchange by opening and closing the valves. The quality of gas exchange is determined by the valve train mechanism on the one hand and by the gas dynamics design on the other hand. The gas dynamics are influenced by the cross sections defined by the valve position and by the valves' opening and closing times. Due to the nature of fuel injection, gas dynamics and gas exchange of internal

combustion engines, some conflicts of goals arise designing such systems. Engineers have to find a compromise in engine characteristics in terms of fuel consumption, torque-progress, maximum power, idle stability, and emissions. To meet demands, additional degrees of freedom are needed. Possible variabilities in valve timing are illustrated in Figure 6.1. Technical implementations are not restricted to a single variable. They also make use of combinations of the presented variations.

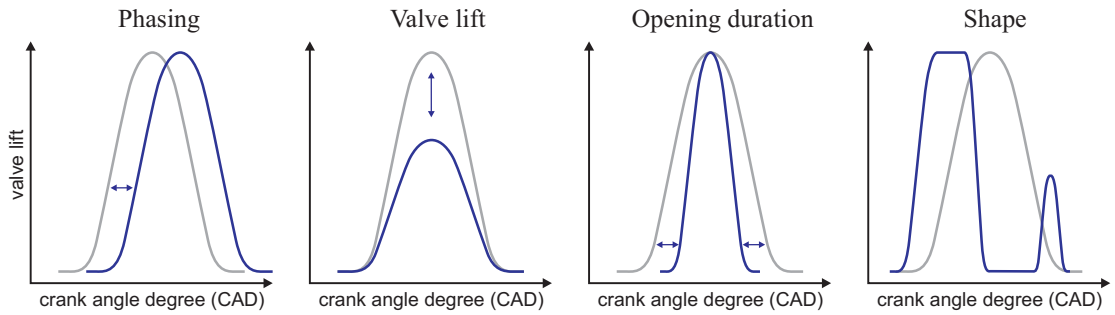


Figure 6.1: Possible valve train variabilities

6.1.1 Types of Variable Valve Trains

In terms of variability, valve trains can be categorized into different groups with increasing degrees of freedom. The first step towards a variable valve train is the application of cam phasing systems, which only alter the phasing of the intake and exhaust lift. By adjusting valve overlapping, exhaust gas recirculation and valve closing times, idling quality, torque curve, emissions and efficiency are improved. Partial dethrottling is possible as the load is controlled by the ratio of exhaust gas recycling. Cam phasers consist of an element that changes the phase of the camshaft relative to the crankshaft by means of rotation. Cam phasers are widely used in series production. They are commonly designed as hydraulic systems based on the functional principle of a slewing motor.

The next generation of variable valve lifts are mechanical systems that modify the valve lift curve. Discretely operating systems called shiftable or switch-over valve lift systems change the shape of the valve lift by switching between up to three different cam contours. The basic idea of such systems is that there are more than one cam contour per valve. Depending on the operating point, switching elements or actuators make sure that the desired cam and corresponding cam follower are engaged. Two stage shiftable systems commonly use a high lift curve for full load and a low lift curve. During partial load, the valve lift on the intake side is reduced and thus the opening time is shortened. As a result, the load control is shifted at least in part from the throttle flap towards the intake valve train reducing pumping losses. Instead of the low lift curve, a zero cam profile can be used for cylinder deactivation. The implementation of a third stage provides for a significantly higher

potential for reducing fuel consumption, since both strategies "low lift" and "cylinder deactivation" can be combined. Alternatively, three-stage systems can be used to apply three different cam profiles for low-end, mid-range, and peak power. Cylinder deactivation can also be realized by modified two-stage systems. For this purpose, transmission elements are used instead of different cam contours to change the valve lift. Switch-off cam followers or switchable hydraulic lash adjusters can generate lost motion to prevent movement of the valve in order to deactivate the valve or cylinder [47].
Shiftable valve train concepts have been installed by car manufacturers since the early 90s. The first high volume system called VTEC (Variable Valve Timing Engine Control) was presented by Honda in 1989 [105], later combined with a cam phasing system [4] and enhanced to a 3-stage system. Afterwards, various discrete operating two-stage systems have been implemented in production engines: MIVEC [92] (Mitsubishi innovative valve timing electronic control system) by Mitsubishi Motors, VVTL-i [183] (Variable Valve Timing and Lift intelligent system) by Toyota, Vario-Cam Plus [28] by Porsche, and AVS [96, 108, 205] (Audi valvelift system) by Audi.
A further adjustment of the gas exchange process within all operating ranges can be accomplished by a continuous adjustment of the valve lift. Mechanical variable valve train concepts commonly use an additional set of intermediate rocker arms or levers between the camshaft and the valve stem as lift scalers. By actuating the intermediate transmission mechanism, the effective relationship of the levers is changed providing the possibility to vary the valve lift continuously between zero and maximum valve lift. In combination with an adaptation of the phase relation, a continuously adjustable valve lift concept permits operation in any map area without the use of the throttle flap. As a result, the combustion process is improved and the gas exchange losses are significantly reduced yielding better fuel efficiency during partial load operation. In addition to fuel consumption, full variable valve lift systems also have the advantage of a clearly improved torque curve and a more direct response.
The only continuously operating mechanical system used in production engines have been introduced in the year 2001 by BMW under the name Valvetronic [58, 198]. In addition, there are several further concepts for mechanically variable valve trains without series implementation.

The next generation of variable valve trains provides additional degrees of freedom by single actuated valves and therefore a higher fuel economy potential. Camless systems allow individual control of each valve by an electromagnetic actuator independently of the crankshaft angle. To minimize the electric power demand, a vibrating armature in combination with springs is commonly used to store the kinetic energy of the valves. At the upper and lower end position, the valve is held by electromagnets. As electromagnetic forces are strongly distance dependent, high currents would be necessary to hold the valves at positions between the end stops. Hence, the valve lift is not varied, but load control is achieved by adjusting the point of intake closing. Compared to mechanical valve trains, where small valve lifts are used to control the load, the additional flow losses at the intake valve are smaller. A further improvement in process efficiency can be achieved by decoupling

load management and generation of in-cylinder gas dynamics. Generation of the in-cylinder gas movement must occur as near to the ignition timing as possible to prevent the charge movement from slowing down. By opening the intake valve once more for a short period, an additional turbulence is excited and yields an improved combustion process.
Further developments of electro-mechanical valve trains also offer the possibility to change the valve lift. For this purpose electric linear motors or rotational actuators may be used. Such systems, along with electro-hydraulic systems, form the fourth generation of variable valve trains. Although various car manufacturers have participated for over twenty years in examining camless engines, vehicles with engines of this type are still not available due to difficulties ensuring reliable performance.

As previously mentioned, the restriction to a constant valve lift results in the fourth generation of variable valve trains, which allow a complete variation of the valve lift curve. Besides electro-magnetic actuators, such systems are mainly driven hydraulically and controlled by fast switching solenoid valves. Applying double acting hydraulic cylinders, there are almost no restrictions to the valve lift curve [37]. But in return, the disadvantages of high energy consumption and high valve closing velocities make series production application difficult. Nevertheless, such camless systems are widely used in combustion research.
Another type of electro-hydraulic valve control systems called UniAir or MultiAir [19], has been launched into series production in 2009 by Fiat. Instead of a double acting piston, the valves are opened by oil pressure and closed by a valve spring. A cam follower, driven by a mechanical camshaft, is connected to the engine valve through a hydraulic chamber. The oil pressure generated by the cam can be controlled by activating hydraulic solenoid valves resulting in a variation of valve timing and maximum valve lift. Furthermore, during one cam revolution multi-lifts are possible. The valve seating velocity is limited by a hydraulic break orifice.

A comprehensive presentation of published variable valve actuation systems is not the objective of this section. For further information and application examples see Basshuysen and Schäfer [199], Dengler [38], Hannibal [90], Hannibal et al. [91], and Scharrer [175].

6.1.2 Benefits of Variable Valve Timing

Variable valve actuation can be used to positively influence specific fuel consumption, emission behaviour, torque characteristics, and maximum power output. A basic goal of varying the valve lifting curves is to lower the charge cycle loss under partial loads by controlling the load solely by varying the valve lift instead of using a throttle. The utilisation of variable valve lift systems increases engine torque in the low speed range and leads to more spontaneous engine response and stable idling. Throttling with the valve has fairly complex effects on the load exchange, combustion, and thermodynamic cycle of the engine. For example, a strategy to decrease pumping losses is to control the load by an early closure of the inlet valve [59]. Löbbert [134] applied this strategy to a naturally aspirated 4-cylinder engine. Due to

a decelerated charge motion resulting in a reduced residual exhaust gas tolerance, no significant improvement in fuel consumption could be obtained in a first step. Only optimisation measures to intensify charge motion like an asymmetric valve lift or masking for a better induction swirl made the potential of a throttle free load control accessible.

This example reveals that a lot of mutually interacting effects have to be taken into consideration when evaluating the influence of variable valve trains on engine behaviour. For a gasoline engine the most important effects concerning fuel efficiency of a variable valve timing system are:

- charge cycling and pumping losses
- flow losses at the valves
- charge motion, mixture generation and in-cylinder flow
- exhaust gas recirculation and residual gas tolerance
- shifting of the load point
- knocking and effective compression ratio
- influence of the VVT system on friction and inertia forces
- energy consumption of the VVT system
- emission behaviour

For a more detailed presentation of the interactions, relations, and implications, refer to the works of Dilthey [42], Löbbert [134], and Scharrer [175] who investigated throttle free load control.

In the last decades, various investigations have been performed to determine to what degree variable valve actuation can improve fuel consumption. But different engine concepts and different boundary conditions (e.g. residual gas content, load condition) make it difficult to compare these results. In addition to variable valve control, series engines are frequently optimised in other ways, making the specific influence of the variable valve actuation difficult to identify by manufacturer's data. Having these difficulties in mind, Figure 6.2 is an attempt to quantify improvements in fuel efficiency by different variable valve actuation systems in the New European Driving Cycle (NEDC). The values are based on a 4-cylinder engine and are adapted from Duesmann [44, 45], and Kirsten [118].

In comparison to spark-ignition engines, the potential for improving diesel engines from variable valve actuation is limited. Since diesel engines are unthrottled, the focus of variable valve timing is placed on emission behaviour by adjusting the exhaust gas recirculation rate and generating charge motion for a better mixture formation [119]. Furthermore, low-end torque can be increased by utilizing exhaust pressure pulsation.

	• **cylinder deactivation** (switchable lost motion element or shiftung cam)	• **cam phaser**	• **cam phaser** • **lost motion element** (switchable tappet, cam follower, ...)	• **cam phaser** • **shifting cam**	• **cam phaser** • **actuated inter-mediate shaft**	• **electro-hydr. system**
						
	intake & outtake	intake & outtake	intake	intake	intake	intake
fuel reduction	approx. 7-8 %	approx. 4 %	approx. 7 %	approx. 8 %	approx. 8 %	approx. 8-15 %
Variability						
- phaising		continuous (intake and outtake)				
- valve lift	discrete: 2 stages standard and zero lift		2-3 stages	2-3 stages	continuous	continuous
- shape			compatible cam contours	independent cam contours		free shape
series production	since 1981	since 1983	since 1989	since 2006	since 2001	since 2009
	Active Fuel Man-agment (GM), ACC (Daimler), VCM (Honda), MDS (Chrysler), CoD (Audi,VW), ...	commonly used by car manu-facturers	VTEC (Honda), MIVEC (Mitsu.), VVTL-i (Toyota), VarioCam Plus (Porsche), ...	Audi valvelift system	Valvetronic (BMW)	MultiAir (Fiat)

Figure 6.2: Different variable valve train systems in series production

6.2 Design and Concept of the Audi Valvelift System

The Audi valvelift system (AVS) manages the valve timing in a gasoline engine in a way that provides greater power and torque in combination with lower fuel consumption. The aim is to open and close the valves in such a way that the correct charge of air is drawn into the cylinders at the right time with minimal losses.
To adjust the valve timing, the Audi 2.8l V6 FSI engine (cp. Figure 6.3) comes with a hydraulic cam phasing system at both the inlet and exhaust camshaft. The hydraulic cam phaser operates like a slewing motor. In combination with a proportional 4/3-way flow control valve, the relative angle between the camshaft and crankshaft can be controlled continuously.

Based on the described cam phasing system, which permits the valve opening and closing times to be varied, the Audi valvelift system goes a step further by varying the valve lift of the intake valves in two stages. The idea of AVS is to use different cam lobes at different operating points. Each intake valve can be actuated by a large and a small cam. For full load, the two inlet valves are opened by the large cam profile (shown in red in Figure 6.4)—ideal for high charge volumes and flow speeds in the combustion chamber. Under partial load, the small cam (green) profiles are used. The valves are opened with a small asymmetric lift that leads to charge air rotating spirally as it flows into the combustion chamber.

As shown in Figure 6.4 the four cam lobes for two valves of one cylinder are mounted on a cam element (2). The cam element is fixed for co-rotation with the base camshaft (1) by an involute gearing but can be moved in an axial direction on the

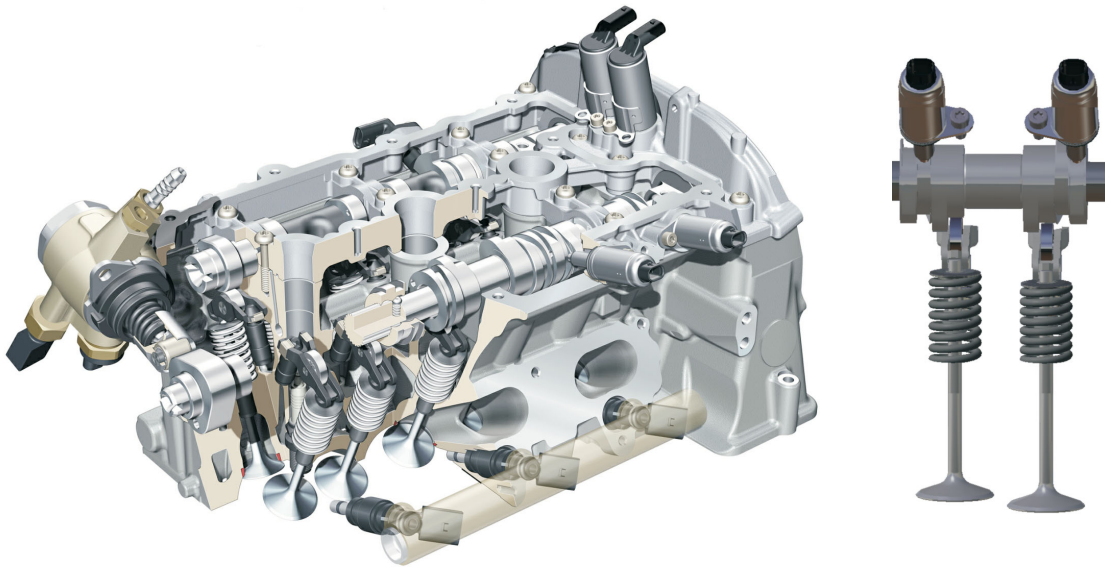


Figure 6.3: Cylinder head with integrated Audi valvelift system

toothed base shaft. Which cam acts on the roller cam follower depends on the axial position of the cam element. The shifting motion is induced by a pair of metal pins, which are extended by bi-stable electromagnetic actuators. The pins are located above an S-shaped control groove milled into the cam element at both ends. When the metal pin is lowered, it engages the spiral groove on the cam element and pushes the cam element in a longitudinal direction. In this way the operating cams are changed from one set to another. An axial movement of the cam element is only possible between two valve actuations. At the end of the switching process the switched pins are in the front end position and are pushed back into the retracted end position by a reset ramp at the end of the spiral grove. Due to electromagnetic induction, a reset voltage is generated within the actuator while the pin slides in. This signal is measured by the engine control unit to confirm successful completion of the switching process. As soon as the cam element has reached its end position, it is locked by a spring-loaded ball stop. The changeover takes place in a range between 700 and 4000 rpm and is completed within one revolution of the camshaft. In order to reverse the cam element, the second actuator is activated.

As a result the gas exchange improves, throttling losses are minimised, and the fuel consumption drops by up to 7% percent. Compared to mechanical systems using intermediate levers to adjust the valve lift continuously, the Audi valvelift system is an easy to manufacture robust system that achieves approximately the same fuel consumption benefit. Dengler [38] compared a two-stage cam shifting system and a fully-variable mechanical valve lift system. In theory, an ideal fully-variable system has a greater potential to minimise throttling losses than a two-stage switching system. Only at the design point of the partial load cam are both systems equal. Having a closer look at technical realisations of fully-variable mechanical systems, some inherent design disadvantages arise: additional force transmitting parts and joints with backlash reduce the stiffness of the mechanism, restrict the slope of the valve lift ramps, and cause higher frictional forces. In addition, the friction level rises

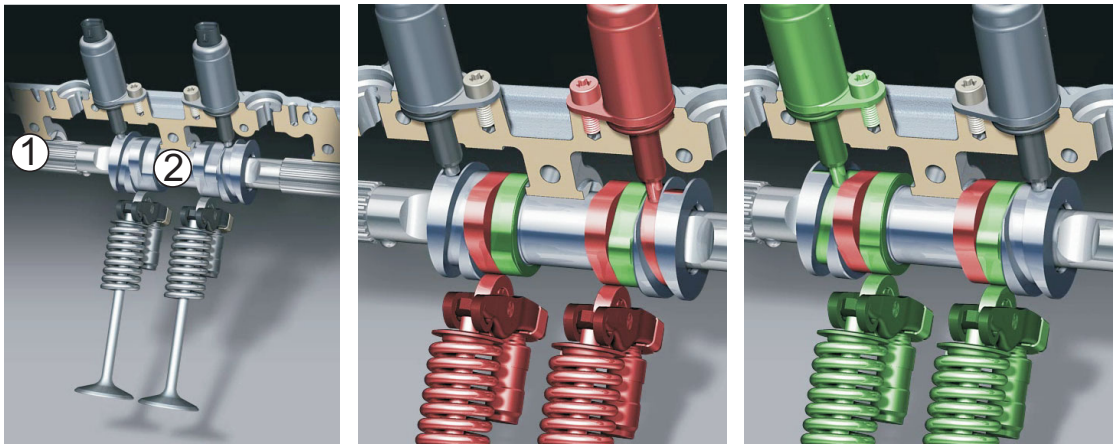


Figure 6.4: Audi valvelift system

as higher moving masses make stiffer valve springs necessary. These are the reasons, amongst others, why a two-stage lift variation comes close to the performance of a fully-variable system, and is even better at operating points for which the shifting system is optimised for. Further comparisons of fully-flexible, two- and three-step variable systems are shown by Sellnau et al. [181].

The Audi valvelift system has been introduced with the naturally aspirated 2.8 litre FSI V6 engine to vary the lift of the inlet valves. Later, the system has also been applied to the exhaust camshaft to improve response and low-end torque characteristic of turbocharged direct-injection gasoline engines [82]. For the 1.8 litre TFSI engine [49,97], the AVS system on the exhaust camshaft is combined with a camshaft adjuster. A similar system is used by Audi and Volkswagen to implement cylinder deactivation as in the 1.4 litre TSI engine [143].

6.3 Valve Train Simulation

In combustion engines valve trains are used to control the combustion process. The valve train is responsible for opening and closing the valves that let fresh air fuel mixture into the cylinder and exhaust gases out. For this purpose, the camshaft rotary motion is converted into a reciprocating motion of the valves. This is almost universally done by using mechanisms that relies on a cam. Usually cams do not act directly on valves since their rotary motion would impose a side force on the end of the valve stem. To avoid lateral loads on the valve stem, cam followers are interposed between the cam and the valve. In direct drive valve trains, linear guided tappets or cam followers are used that rest upon the cams. Direct drive offers very good stiffness of the valve actuating system and relatively modest masses in motion. Indirect drive valve trains include systems with rockers or levers rotating upon fixed pivots.
In modern passenger car engines a trend towards roller-actuated cam followers with hydraulic lash adjustment can be observed. The use of rolling bearings has the effect of friction reduction, particularly in the lower speed range. This reduction in friction is paid for with a significant reduction in damping of torsional vibrations at the camshaft, which has negative effects on the timing chain. In comparison with sliding cam follower concepts, these systems are heavier. Consequently, they are not used in sports car engines because in high-revving engines, low moving masses and high system stiffness are basic prerequisites.

6.3.1 Literature Survey

When designing mechanical valve train systems there is generally a conflict between steep valve lift curves for an ideal gas exchange and low valve spring forces for reducing friction. To optimise the design within these oppositional requirements, several aspects including moving masses, cam contours, and dynamic characteristics of components such as valve springs and hydraulic lash adjusters have to be considered. Hence, a lot of research has been done in the field of valve train dynamics.
The first simulations were carried out in 1956 on electronic analogue computers by Franke [66]. Hofbauer used a dynamic valve train model to optimise the cam lobe [102]. Pisano and Freudenstein derived a dynamic model of a high-speed cam follower system with a distributed parameter approach for the return spring and verified it with the aid of experimental data [165]. Further refinement was done by Hanachi and Freudenstein by considering energy dissipation, hydraulic lifters and damping due to nested valve springs. By identifying the predominance of Coulomb damping at high speeds, the predictive capability of the dynamic model could be improved [89].
Later on Phlips and Schamel also considered the interactions between valve lift and gas exchange processes [164]. Kreuter and Pischinger [122] investigated the influence of oil film effects and lubrication on the effective values of damping and stiffness within valve trains. Degenhardt investigated friction conditions within

valve trains using theoretical approaches and experimental studies [35]. The interactions between crankshaft and valve train, as well as the influence of shaft elasticity in a complete timing assembly have been analysed by Fritzer [72]. Furthermore, he compared analytical approximation methods and numerical integration for the analysis of the vibrational behaviour.
Kalser et al. [117] used valve train models to reduce the valve train noise in multi-valve engines by optimising the cam profile. Ross [172] extended dynamic valve train modelling by bending and torsional modes of the camshaft. Brands [24] presents the integration of elastic valve train components and elastohydrodynamic methods in multi-body environments for valve train simulation. Different levels of detail of valve train and timing assembly models were used by Mendl [141] depending on the objective of the simulation (e.g. strength, dynamics, acoustics, or gas exchange processes).
Engelhardt [50] applied the methods of non-smooth mechanics to valve train simulation. Förg [65] et al. compared different contact models within the valve train. Focus has been directed towards contact between the cam and the cam follower. Huber et al. used non-smooth methods in hydraulics for modelling the hydraulic lash adjuster [109]. Krüger [123] investigated the interactions between valve actuation and the hydraulic cam phasing system.

Simulation has become a standard method for virtual prototyping in the valve train design process. Hence, libraries for valve train simulation at different levels of detail are included in most established multi-body system simulation tools; For example FEV Virtual Engine, ADAMS/Engine, LMS DADS, and SIMPACK Engine. However, these proprietary tools are restricted to smooth mechanics and continuous systems and do not offer the possibility of implementing new systems or components easily.

6.3.2 Valve Train Modelling

The AVS valve train investigated in this work is based on an end-pivot roller rocker arm valve train with overhead camshaft. The exhaust valve train consists of the roller pivot-mounted in the rocker arm (1), the hydraulic lash adjuster or HLA (2), the valve spring (3), the valve (4) and the cam (5) (cp. Figure 6.5). To change the engaged cam contour at the intake side, the basic valve train is extended by the AVS cam element (6) and the electromagnetic actuator (7). The distance between end pivot, roller and valve stem defines a transmission ratio between cam lift and valve lift enabling small cams.

The valve train is modelled using rigid bodies for the roller and the rocker arm. The valve and valve spring are implemented as flexible bodies. Planar motion is considered for a single valve actuation subunit. To describe the motion during switching operations, the cam element is extended by an additional degree of freedom in the axial direction. Detailed models and the internal dynamics of the valve spring, the hydraulic lash adjuster, the cam element with ball stop, and the actuator are presented in the next chapters.

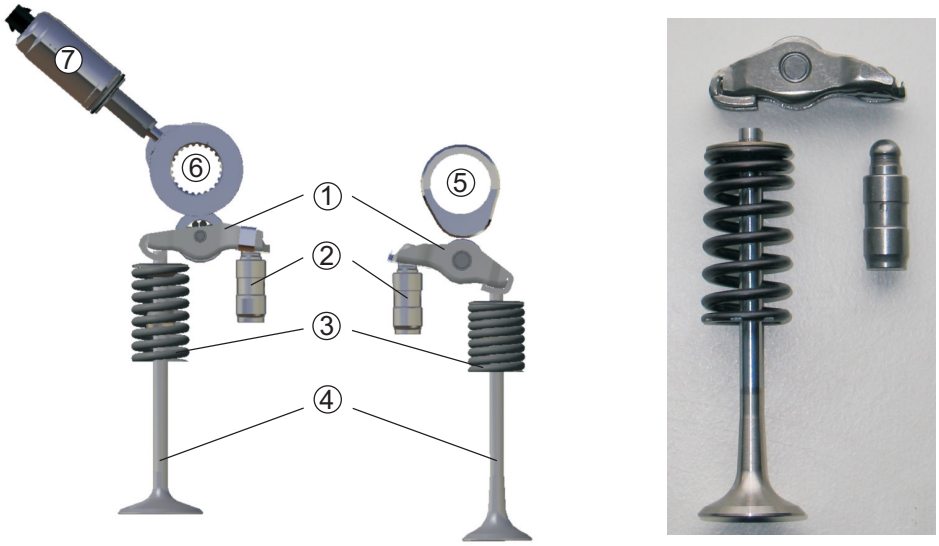


Figure 6.5: Intake and exhaust valve actuation system

The single components are connected via four unilateral contacts: HLA to rocker arm (circle to circle), rocker arm to valve stem (circle to line), valve to valve seat (point to line), and cam to roller (circle to parametrised curve), as well as one bilateral connection or joint with one rotational degree of freedom between roller and rocker arm.
The roller follows the cam contour due to a frictional contact between these two bodies. For the exhaust planar and for the intake spatial Coulomb friction is supposed. The camshaft rotary motion is converted into a reciprocating motion of the valves by the rocker arm. The rocker arm actuates the valves via a frictional contact. Due to the dynamics of the system, impacts can occur between the cam and the roller, between the rocker arm and the valve, between the rocker arm and the HLA, as well as in the valve seat. Coil collisions in the valve spring or the internal dynamics of the HLA are further sources of discontinuities.

Due to the multitude of different contacts and hydraulic sub-models with nearly incompressible oil volumes, valve trains are suitable systems for applying non-smooth mechanics. Stiff differential equations are avoided by treating contacts as rigid and oil volumes as incompressible.
Within the derived model, all contacts and oil volumes can be considered either as rigid using set-valued force laws or as elastic using single-valued force laws. By using a completely rigid model, one could see pulsating contact forces leading to transitions between open and closed contacts at a very fast time scale, especially at the valve's closing position. Therefore, at least the contact between the rocker arm and the valve should be modelled elastically.

6.3.3 Contact Kinematics

To describe the contact between roller (index 1) and cam (index 2), the potential contact points $\boldsymbol{r}_{C_i}$ on the contours need to be determined, as well as the gap distance g_N and the relative velocities $\dot{g}_N$ and $\dot{\boldsymbol{g}}_T$ (see Figure 6.6). This information is used for evaluating the force laws, as well as for calculating the generalised force direction $\boldsymbol{W}$.

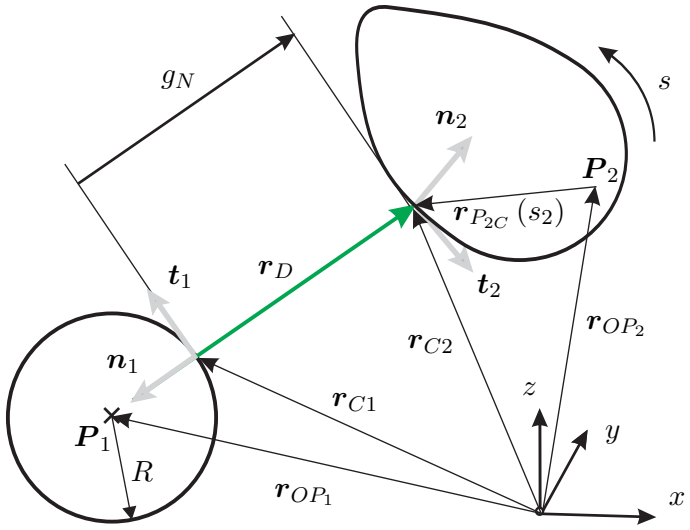


Figure 6.6: Kinematics of the contact between cam and roller

To obtain smooth contact forces if the roller and cam roll off one another without impacts, periodic cubic splines for each coordinate are used to interpolate the cam geometry $\boldsymbol{r}_{C_2}$. A single point on the contour can be referenced by the scalar contour parameter s. As the cam contour is plane the bi-normal direction is fixed and parallel to the camshaft. The tangential and normal vectors $\boldsymbol{t}_2$, $\boldsymbol{b}_2$, and $\boldsymbol{n}_2$ are obtained using standard differential geometry methods [27]. The modified distance vector $\boldsymbol{r}_{\hat{D}}$ is defined by:

$$\boldsymbol{r}_{\hat{D}}(s) = \boldsymbol{r}_{C_2}(s) - \boldsymbol{r}_{OP_1}\,, \tag{6.1}$$

whereas $\boldsymbol{r}_{OP_1}$ is the centre of the roller with radius R. The contour parameter for potential contacts can be calculated by solving

$$f(\boldsymbol{q},s) = \boldsymbol{r}_{\hat{D}}^T\,\boldsymbol{t}_2 = 0 \tag{6.2}$$

with Newton's method or with the Regula Falsi method in the case of a non-converging solution. By minimising the distance of the surfaces $g_N = \boldsymbol{r}_{\hat{D}}T\,\boldsymbol{n}_2 - R$ the actual solution s is obtained. The differential geometrical values of body 1 (roller) are in accordance with Newton's third law of reciprocal actions given by:

$$\boldsymbol{n}_1 = -n_2 \tag{6.3}$$

$$\boldsymbol{t}_1 = -t_2 \tag{6.4}$$

$$\boldsymbol{b}_1 = -b_2 \tag{6.5}$$

$$\boldsymbol{r}_{C_2} = \boldsymbol{r}_{OP_1} - R\,\boldsymbol{n}_1\,. \tag{6.6}$$

The relative contact velocities and the generalised force directions can be calculated following Eq. (2.35) and (2.36). At the intake side, there are two different cam contours for one roller. Hence, the active cam contour parametrisation is chosen depending on the axial position of the cam element.

7 AVS Cam Element

This chapter deals with the cam element including a spring loaded ball stop and toothing. The simulation model is derived and validated with experimental data obtained from a component test rig. The test rig is also used for the identification of friction parameters.

7.1 Modelling the Cam Shifting System

The cam element is axially displaceably mounted on a toothed camshaft. Depending on the position of the cam element the smaller or bigger cam profile is active. Since errors during the changeover can cause engine damage, the cam element must be seen as a critical component and will therefore be examined more closely.

7.1.1 Contact Kinematics

The lateral contact between the pin and the spiral groove is modelled as a contact between a cylinder and a spatial curve. The geometry of the lateral flank of the groove is parametrised by the scalar contour parameter s, which can be interpreted as the angle between the beginning of the groove and the actual position in cylindrical coordinates (cp. Figure 7.1).

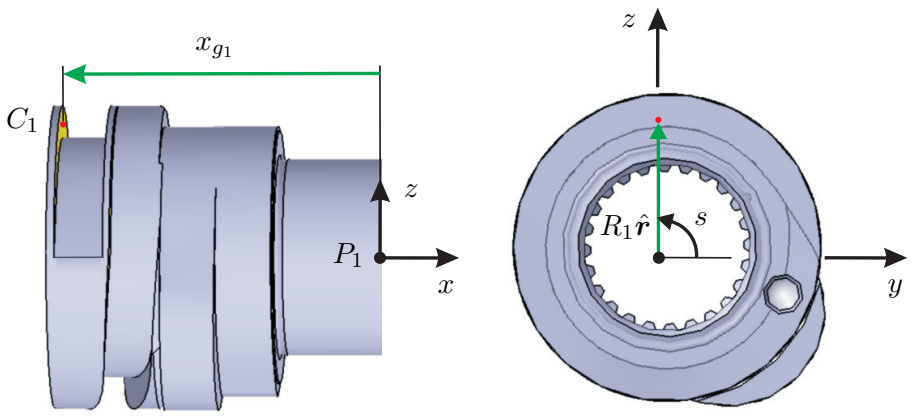


Figure 7.1: Parametrisation of spiral groove

A single point C_1 of the contour is given by:

$$\boldsymbol{r}_{C_1}(\boldsymbol{q}, s) = \boldsymbol{r}_{P_1}(\boldsymbol{q}) + R_1\,\hat{\boldsymbol{r}}(s) + \hat{\boldsymbol{e}}_x(\boldsymbol{q}) x_{g_1}(s) \tag{7.1}$$

with radial direction $\hat{\boldsymbol{r}}^T = [0, \cos(s), \sin(s)]$ and axial direction $\hat{\boldsymbol{e}}_x$. The distance in axial direction x_{g_1} is approximated by cubic spline interpolation. The planar projection of the contact kinematics is drawn in Figure 7.2.

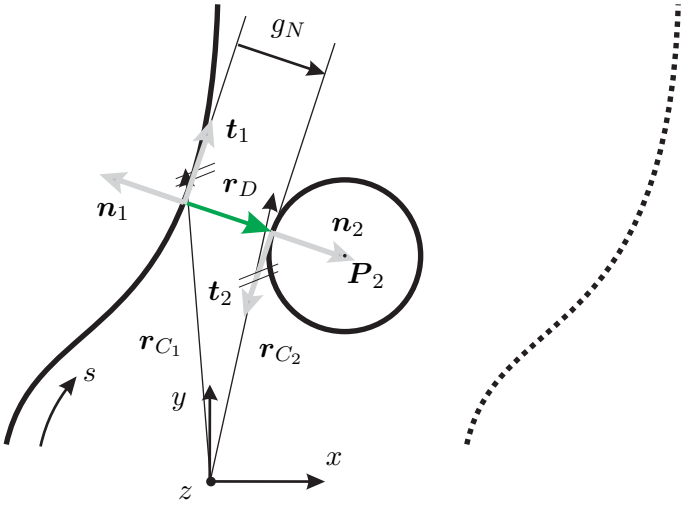


Figure 7.2: Planar projection of the contact kinematics pin to groove

In analogy to the cam roller contact, the defining equation for potential contact points is given by:

$$f(\boldsymbol{q},s) = \boldsymbol{r}_D^T\,\boldsymbol{t}_1 = 0\,. \tag{7.2}$$

In contrast to the previous contact problem, both contours are not within the same plane. Hence, the solving of Eq. (7.2) is adapted to consider the spatial contact situation. Equation (7.2) is solved in two steps: First, using cylindrical coordinates the groove contour is approximated by a helical curve with constant slope. Thus, the contour can be described using analytical functions and an analytical approximate solution s_0 exists. Second, this approximation is used as a starting value for a fixed point iteration to calculate the exact value of the contour parameter. The computational procedure of the fixed point iteration is obtained by an orthogonal projection of the centre axis of the pin in the tangent space of the contour:

$$\boldsymbol{r}_{C_1}^{i+1} = \boldsymbol{r}_{P_2} - \boldsymbol{n}\,\boldsymbol{n}^T \left(\boldsymbol{r}_{P_2} - \boldsymbol{r}_{C_1}^{i}\right)\,. \tag{7.3}$$

With s being a solution of Eq. (7.2), the remaining kinematic quantities are calculated according to Section 6.3.3. The vertical contact problem between the ground of the spiral groove (reset ramp) and the end surface of the pin is also solved with the methods presented for the cam roller contact.

7.1.2 Toothing

The cam element is fixed for co-rotation with the basic camshaft by an involute gearing but can be moved in the axial direction on the spline shaft. The gearing with 24 cogs is designed to bear the load along its flanks. The level of detail for the gearing depends on the intention of the simulation. For stationary simulations with kinematically excited camshaft at constant speed, it is sufficient to connect the relevant degrees of freedom of camshaft and cam element by a bilateral constraint. If alternating loads are applied to the camshaft e.g. due to combustion forces or due to an active cam phasing system, an appropriate model for the toothing is a frictional unilateral contact with backlash.

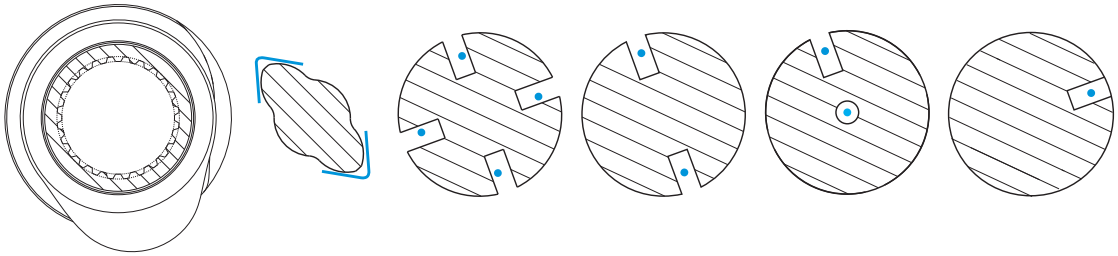


Figure 7.3: Possibilities for reducing spline shaft contact geometry

To achieve an efficient model and avoid multiple overdetermined systems, the number of possible contact points must be reduced. Possibilities to reduce the spline shaft contact geometry are illustrated in Figure 7.3. Two different levels of detail are used. A direct implementation of the spline shaft geometry is to approximate the real contour by ring segments or cubic splines. Two teeth are sufficient to constrain the planar motion. In most applications it is possible to use contacts between points and planes. For camshafts and cam elements with only one relative rotational degree of freedom, one plane to point contact is used to model the backlash in the circumferential direction. The ball of the ball stop has one degree of freedom relative to the camshaft. Numerical studies reveal that the results of all approaches do not differ significantly. The use of contact pairings based on points and planes shows a slightly reduced computational effort.

7.1.3 Ball Stop

The movement of the cam element is mainly affected by three forces. The contact force between the spiral groove and the pin accelerates the cam element during the changeover and ensures that the changeover takes place during the base-circle phase. The resultant force from the ball stop decelerates the cam element at the beginning, accelerates it at the end of the changeover, and provides a static contact force to fix the cam element in the end positions against the end stop. Lastly, frictional forces acting against the direction of movement must be taken into account.

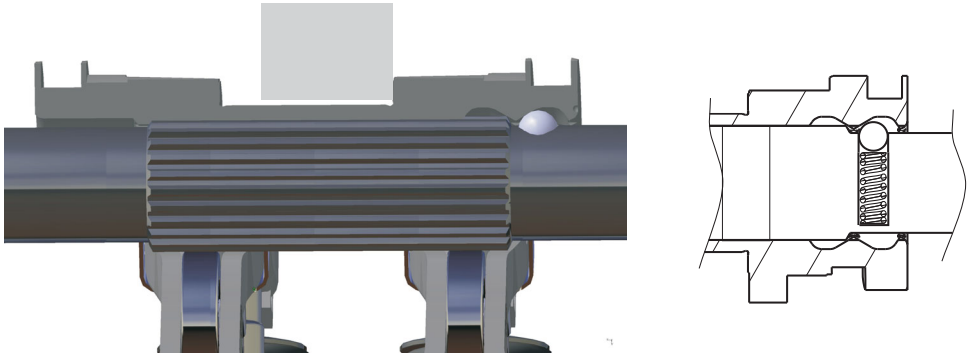
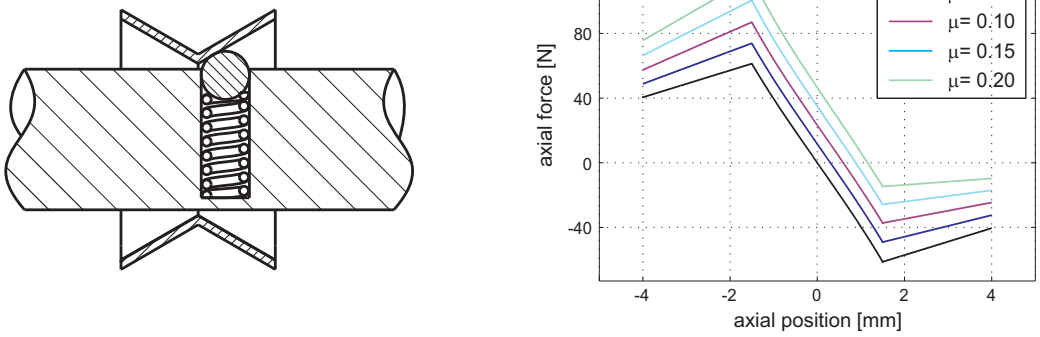


Figure 7.4: Cam element with ball stop

The contact between ball and cam element is described by a unilateral contact with friction between a circle and a cone. The geometry of this contact pairing is outlined in Figure 7.5(a). To make this type of contact problem (Eq. (2.33)) analytically solvable, the apex of the cone is replaced by an infinite segment of a circle to ensure parametric continuity. The simulation results of the axial forces necessary to push the cam element against the ball stop in axial direction are shown in Figure 7.5(b) for Coulomb friction with altering friction coefficients. As the axial force is very sensitive to the friction coefficient, the friction force is experimentally investigated in the next section.



(a) Geometry ball stop

(b) Axial forces ball stop

Figure 7.5: Simulation of axial forces with different Coulomb friction coefficients

7.2 Test Rig for Friction Measurements

Friction is a complex phenomenon and difficult to determine. Hence, in technical simulations friction is often neglected or roughly estimated. To get reliable friction models and parameters, a test rig for both static and dynamic measurements has been developed. The test rig is designed to quantify friction forces at the spline shaft. The forces due to friction are measured for different relative velocities and

lubrication conditions. Measurements have been performed with and without ball stop.

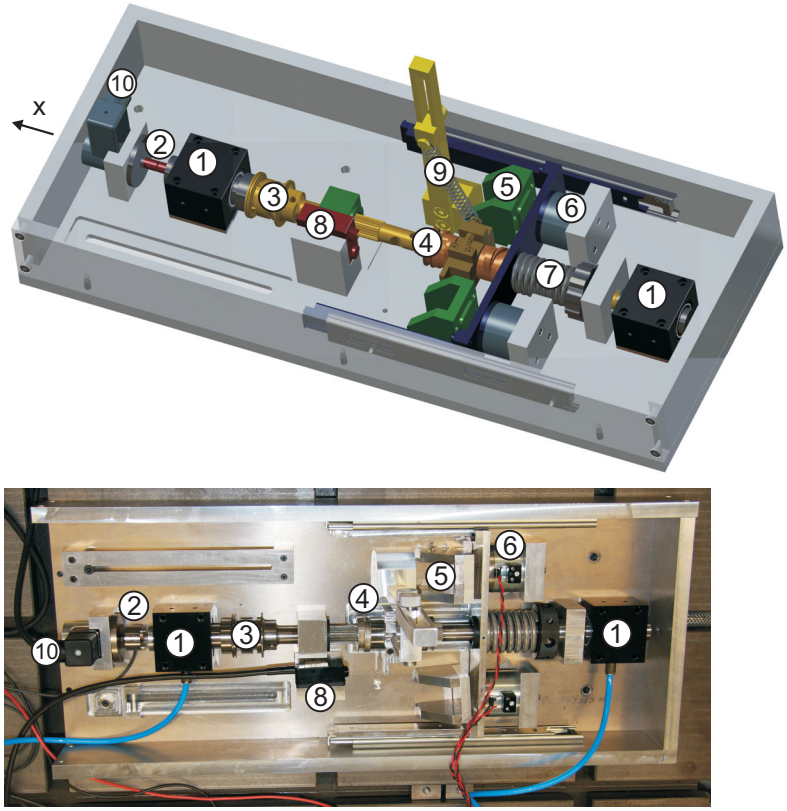
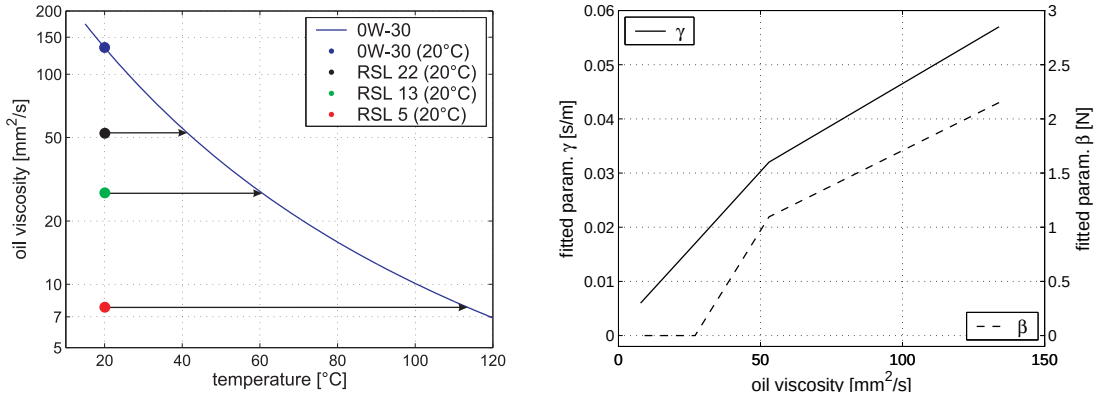


Figure 7.6: Test rig for friction measurement

The experimental setup is shown in Figure 7.6. The camshaft (3) is mounted with almost frictionless air-lubricated bearings (1). Using a compressed spring (7) releasable by two electromagnets (6), the cam element (4) is accelerated in an axial direction. The section for accelerating the cam element is limited by two end stops (5). After the acceleration phase, the position of the cam element is gauged by a laser sensor (8). Assuming zero friction at the air bearings, the friction force between cam element and camshaft can be measured with a force sensor (2) at the end of the camshaft. For different normal forces at the friction contact, the prestress of the spring (9) is varied. To fix the camshaft in axial direction without interlocking and without applying radial bearing forces, an electromagnet (10) is activated after the camshaft is lifted up by the air bearings. A more detailed description can be found in [53].

The influence of temperature is determined indirectly using oils with different viscosities at room temperature as it is very difficult to operate the test rig, air-lubricated bearings, and sensors at high temperatures. The underlying engine oil is Castrol TOPUP 0W-30. Instead of heating up the engine oil, the hydraulic fluid AVIA RSL is used. It is similar to mineral oil and different viscosity grades are available. The viscosity values and corresponding temperatures are illustrated in Figure 7.7(a).



(a) Oil viscosities used in experiment (b) Fitted parameters as function of viscosity

Figure 7.7: Viscosities of used oils and resulting viscosity dependency of fitted parameters

7.3 Results

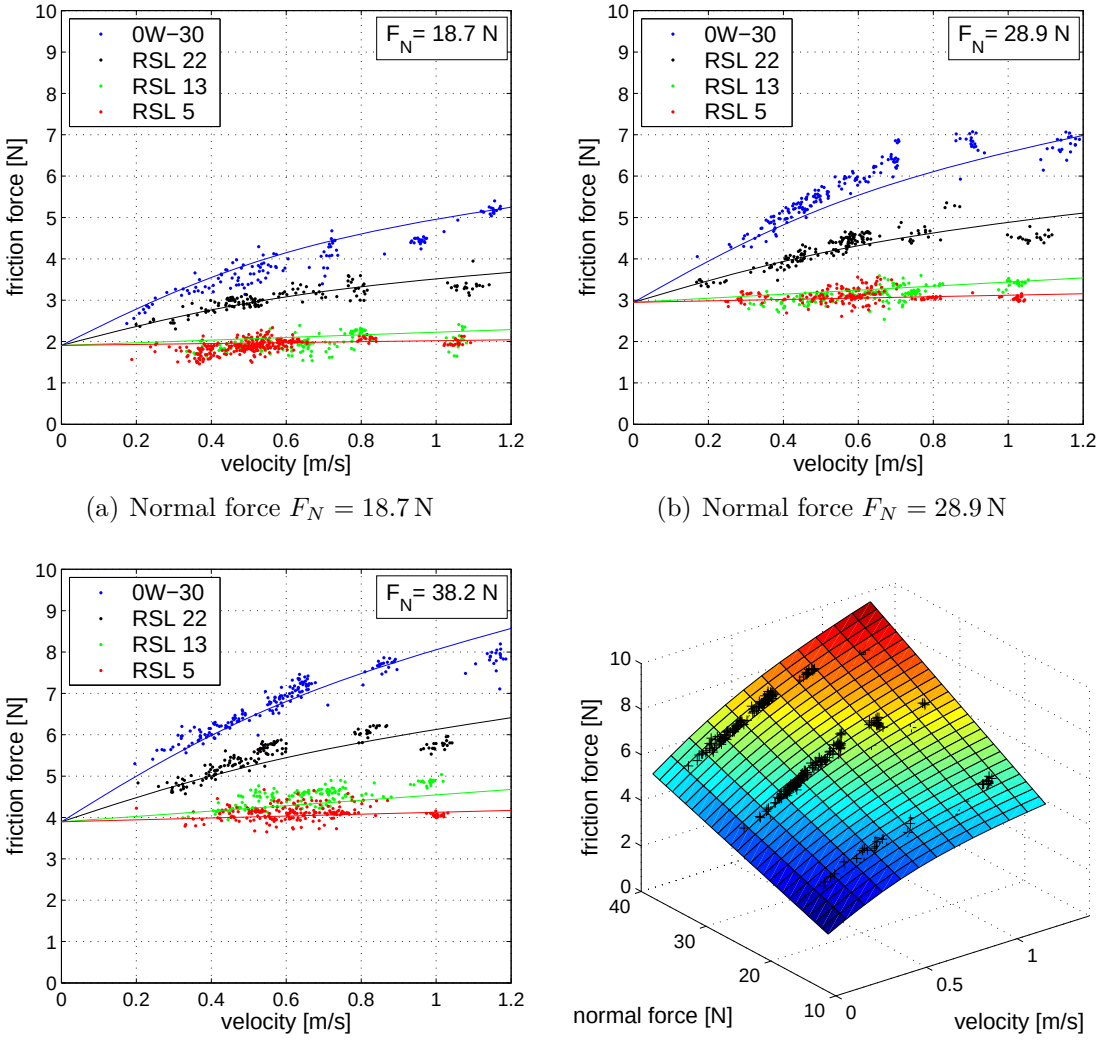
7.3.1 Friction Model

In Figure 7.8 the measured friction values for different oil viscosities are plotted as dots against the relative velocity. Although the data shows high variation, some trends can be identified.

Higher normal forces F_{normal}, growing velocities v, and rising oil viscosity have the effect of increasing friction forces F_{friction}. These results suggest that with increasing velocity hydrodynamic effects become more and more significant. Interpolating the test series to zero velocity yields an oil-independent friction force. Relating the interpolated friction forces with the normal forces a constant friction coefficient can be identified. Analysing the data with curve fitting methods, the following friction model has been derived:

$$F_{\text{friction}} = \alpha\, F_{\text{normal}} + \beta\, v + \gamma\, v\, F_{\text{normal}}. \tag{7.4}$$

The lines of best fit for different normal forces and different oil viscosities are shown in Figures 7.8(a) to 7.8(c). The surface of best fit compared with the individually measured data points is drawn exemplary for Castrol TOPUP 0W-30 in Figure 7.8(d).
The parameters α, β, γ for the derived friction model are calculated by curve fitting supposing that the parameter α is constant for all test series. The results are shown in Table 7.3.1. The viscosity dependency of β and γ is illustrated in Figure 7.7(b). For small viscosities both parameters are approaching zero. For high velocities, the friction force seems to become independent from the velocity and a non-linear velocity dependency would be helpful in this region.



(a) Normal force $F_N = 18.7\,\mathrm{N}$

(b) Normal force $F_N = 28.9\,\mathrm{N}$

(c) Normal force $F_N = 38.2\,\mathrm{N}$

(d) Castrol TOPUP 0W-30 (20° C)

Figure 7.8: Friction characteristic

Table 7.1: Best fit parameters

oil	$\nu\ [mm^2/s]$ (20° C)	$\alpha\,[-]$	$\beta\,[N]$	$\gamma\,[s/m]$	RMSE [N]
0W-30	134	0.102	2.15	0.057	0.34 N
RSL 22	53	0.102	1.09	0.032	0.23 N
RSL 13	27	0.102	0.0	0.017	0.23 N
RSL 5	7.8	0.102	0.0	0.006	0.17 N

7.3.2 Axial Forces due to Friction and Ball Stop

The derived friction model (Eq. (7.4)) is used in simulations to approximate friction forces at the contacts between camshaft and cam element, as well as between cam element and ball stop. In Figure 7.9 the axial forces at the camshaft during the movement of the cam element across the ball stop are shown. Simulated and

experimental results agree very closely. The dotted lines are obtained under the assumption of dry Coulomb friction with coefficient $\mu = 0.198$. The solid lines are obtained using the derived friction characteristic for a lubricated contact.

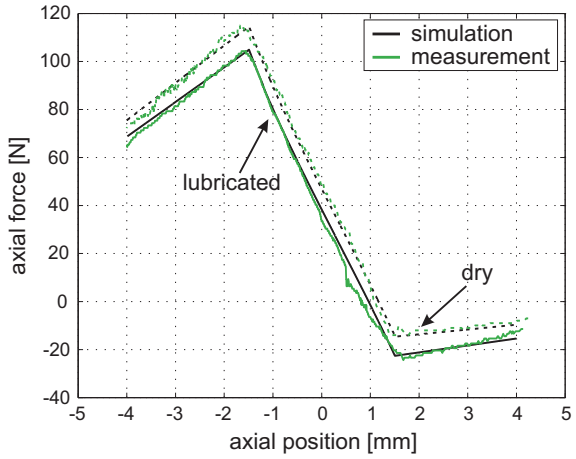


Figure 7.9: Ball stop - experiment and simulation

8 The Electromagnetic Actuator

This section deals with the actuator for driving the AVS cam element. Starting with the geometrical and technical configuration, a physical simulation model based on Maxwell's equations is derived and compared to experimental results.
In normal operation the system response is very fast and the pin is fixed at its end stop before the changeover begins. Hence, the actuator has only a slight influence on the overall system dynamics. The requirements of the actuator model are that it can be easily integrated into the multi-body formalism, it allows for short computational times, and it predicts the magnetic force and thus the movement of the pin with adequate accuracy. The actuator model is not intended to be used for the electromagnetic design and optimisation of the actuator, or to calculate detailed magnetic fields. Hence, an analytical model for a simplified geometrical topology is the best choice in this work instead of a detailed FEM model.

8.1 Types of Actuators

As the rotation of the camshaft is the source of energy for the switching operation, the force delivered by the actuator is not used to drive the cam element directly and therefore can be relatively small. It must be only large enough to overcome friction and the inertia forces of the moving parts.

Actuators can be distinguished by the type of their energy source, as illustrated in Figure 8.1. A comparison of different actuator concepts can be found in [112] and [196]. Magnetic actuators are ideal for this kind of application as pneumatic or hydraulic systems need a lot of periphery and the operating displacement of piezo or electrostatic actuators is very small.

In principle, all kinds of magnetism and corresponding phenomena can be described by the theory of quantum mechanics [137]. The concept of elementary particles having a spin in combination with the Pauli Exclusion Principle for fermions is the fundamental quantum theoretical basis for most magnetic properties of solids. As this theory is not easily explained and not practical for technical problems, a more phenomenological approach to magnetism is often used [115]. In a technical context magnetic actuators can be classified by the technical or physical principles used for generating the magnetic force:

- reluctance force
- electrodynamic Lorentz force

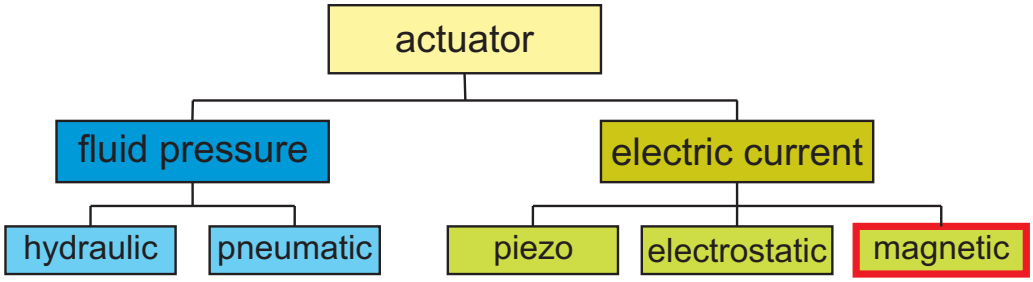


Figure 8.1: Energy sources of actuators

- magnetostriction
- magnetic shape-memory alloys
- ferro-fluids and magneto-rheological fluids

8.1.1 Reluctance Actuators

The principle of magnetic reluctance (reluctance = magnetic resistance) describes magnetic forces between materials with different magnetic permeability. Stepwise changes in the permeability μ_r induce magnetic forces at the corresponding bounding surfaces. These reluctant forces try to minimise the overall reluctance of the system.
Similar to an electric field in which an electric current follows the path of least resistance, a magnetic field results from the magnetic flux following the path of least magnetic reluctance. Magnetic flux always forms a closed loop, but the path of the loop depends on the reluctance of the surrounding materials and is always concentrated around the path of least reluctance. The concentration of flux in low-reluctance materials such as iron forms strong temporary poles and causes mechanical forces that tend to move the materials towards regions of higher flux and to minimise the magnetic resistance.
In contrast to electrodynamic Lorentz forces, the direction of reluctant forces between ferromagnetic materials is independent of the field direction. Reluctant forces are always attractive and try to minimise the air gap. A second magnet is needed to generate repulsive forces.

The properties of an electromagnetic reluctance actuator can be expanded if, in addition to the coil excitation, a permanent magnet as a complementary source of the magnetic field exists. This arrangement is called a bipolar or polarised electromagnetic actuator [171]. As magnetic fields and forces are active even without coil current, powerless clamp forces can be realised. Furthermore, polarised actuators have a very good force-to-weight ratio and a high energy and force density.

The principle of reluctance is used in reluctant motors, electromagnets (magnetic clamp, lift, switch, or proportional solenoid), and high power shakers [99] or actuators. Beside the present AVS switch actuator, reluctant actuators, mainly proportional solenoids, have a wide range of automotive applications as an alternative to

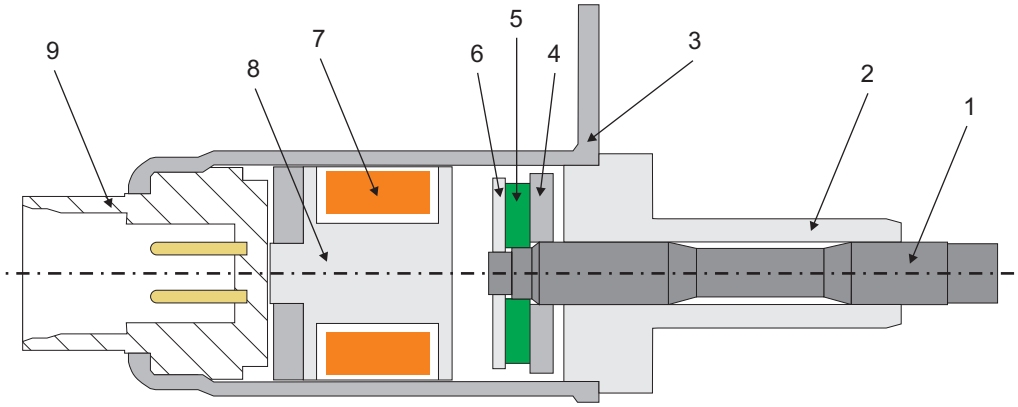


Figure 8.2: Drawing of the electromagnetic actor

expensive servo-hydraulic valves from aerospace engineering [133]. They are used within direct-shift gearboxes, cam phasing systems, and fuel injection systems.

8.1.2 Non-Reluctance Actuator Classes

Actuators or systems based on the electrodynamic principle utilise the so called Lorentz force, which acts on current-carrying wires in magnetic fields. Unlike reluctant forces, the direction can be attractive or repulsive depending on the field and current direction. Although the force density of such actuators is smaller, this principle is widely used for electric motors and generators.

Magnetostriction is the effect of altering the magnetisation of a material that results in a change of its shape [17, 129]. Magnetostrictive materials can convert magnetic energy into kinetic energy, or reverse, and are used to build sensors or actuators with a small operating displacement (e.g. ultrasonic or audio devices). Magnetostriction is also a source of the electric hum around transformers.

Magnetic shape-memory alloys and magneto-rheological fluids are used within special applications such as aeronautics and medical engineering. Please refer to literature [113, 208] for further information.

8.2 Design and Concept

The actuator used for the Audi valvelift system is a polarised electromagnetic actuator. To induce the motion of the cam element, the metal pin of this bistable reluctance actuator is extended.

The functional principle of this actuator is the repulsion force of two magnets that are magnetised in opposite directions. As painted in Figure 8.2, one of these magnets is a solenoid consisting of a coil (7) and a ferromagnetic core (8). The second magnet is a permanent magnet (5) amplified by a pole disk (6) and an armature (4). These

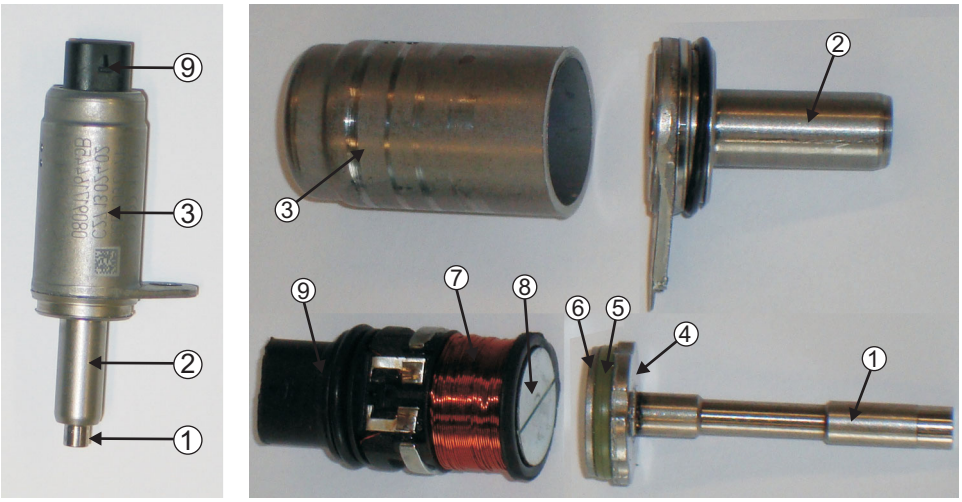


Figure 8.3: Components of the electromagnetic actor

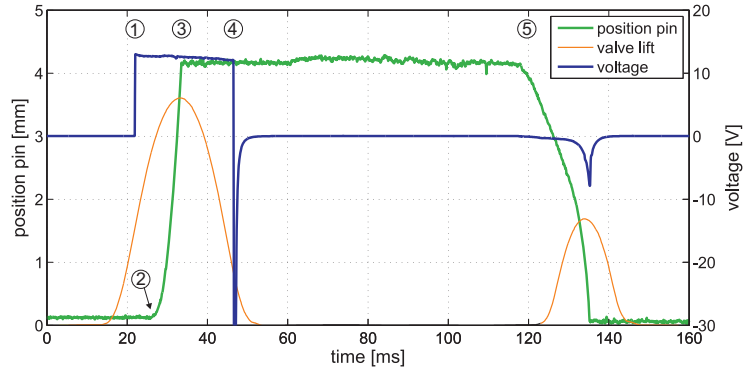
Table 8.1: Components of the electromagnetic actuator

Pos.	Description	Material
1	pin	steal
2	guide tube	steal
3	box	steal
4	armature	steal
5	permanent magnet	steal
6	pole disk	steal
7	solenoid coil	copper
8	coil core	steal
9	connector	plastic

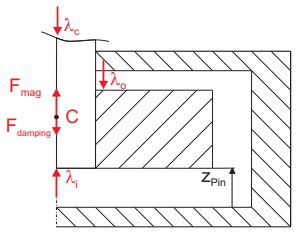
three parts are attached to the pin (1) and compose the moving assembly of the magnetic actuator (see Figure 8.3 and Table 8.2). Due to the permanent magnet, the metal pin is held at both end stops with a defined magnetic force. Electric energy is only needed to extend the pin. As the pin is pushed back by the spiral groove of the cam element, no magnetic forces are supplied that could retard the pin. Hence, the actuator can be designed as a unidirectional device.

In Figure 8.4(a) the operational sequence of a switching process from the large to the small cam element is illustrated. Data is recorded at a crankshaft speed of 1000 rpm [192].

Before the switching operation starts, the iron core of the solenoid is magnetised by the permanent magnet to yield a static magnetic force that holds the pin at its inner end stop. To start the changeover, a 12V voltage step is applied to the coil (point 1). The inductance L of the coil causes a delay in current rise, and consequently in the rise of the magnetic field. Hence, the pin reacts with a time delay on the voltage step and starts to move (point 2) from the inner to the outer end stop (point 3). At point 4 the voltage is shut down and the magnetic field energy induces a



(a) Switching process from large to small cam at 1000 rpm: supply voltage, pin position and valve lift



(b) Mechanical Model: Moving parts and end stops

Figure 8.4: Illustration of switching process and mechanical Model

voltage overshoot. After changing the cam contour, the pin is pushed back by the contact forces between the tip of the pin and the bottom of the spiral grove. As the permanent magnet is attached to the moving part of the actuator, the magnetic flux in the coil changes and induces a voltage signal (point 5). As this signal is used for monitoring the changeover, induction due to pin motion has to be considered in the actuator model. A simulation model coupling both physical domains is necessary.

8.3 Physical Model of the Electromagnetic Actuator

Both physical domains—mechanics and electrodynamics including magnetostatics—are modelled to achieve a dynamic model of this mechatronic component. The mechanical model is shown in Figure 8.4(b). The motion of the pin is constrained by the inner and outer end stops as well as by the contact between the spiral groove and the pin yielding the constrained forces λ_i, λ_o and λ_c. With known magnetic force $\boldsymbol{F}_{\text{mag}}$, the equations of motion can be derived with the methods described in Section 2.1. As the pin is built similarly to a pneumatic piston, it has to compress or eject air. This effect is most notable near the end stops. Therefore, a modified viscous damping force $\boldsymbol{F}_{\text{damping}}$ with quadratic velocity dependence is used.
Calculations of the magnetic forces as well as induction and coupling between mechanics and electrodynamics, are presented in the next section.

8.3.1 Electromagnetic Field Theory

This section presents a summary of electrodynamics and field theory [56, 131]. The basic equations in electrodynamics are a set of partial differential equations postulated by Maxwell [139] that combine electric field quantities (electric field $\boldsymbol{E}$, electric

displacement field $\boldsymbol{D}$) with magnetic field quantities (magnetic field strength or magnetising field $\boldsymbol{H}$, magnetic induction or flux density $\boldsymbol{B}$):

$$\operatorname{rot} \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t} \qquad \text{(Farady's law)} \tag{8.1a}$$

$$\operatorname{div} \boldsymbol{D} = \rho \qquad \text{(Gauss's law)} \tag{8.1b}$$

$$\operatorname{rot} \boldsymbol{H} = \boldsymbol{j} + \frac{\partial \boldsymbol{D}}{\partial t} \qquad \text{(Ampere-Maxwell's law)} \tag{8.1c}$$

$$\operatorname{div} \boldsymbol{B} = 0 \qquad \text{(Gauss's law for magnetism)}\,. \tag{8.1d}$$

Electric fields are generated by a charge density ρ or electromagnetic induction $\dot{\boldsymbol{B}}$. The source of magnetic fields are current densities $\boldsymbol{j}$ or displacement currents $\dot{\boldsymbol{D}}$. The field quantities are linked by the material properties permittivity ε (ε_0, ε_r) and permeability μ (μ_0, μ_r) [142].

$$\boldsymbol{D} = \varepsilon_0 \varepsilon_r \boldsymbol{E} \doteq \varepsilon \boldsymbol{E} \tag{8.2a}$$

$$\boldsymbol{B} = \mu_0 \mu_r \boldsymbol{H} \doteq \mu \boldsymbol{H} \tag{8.2b}$$

Together with Maxwell's equations (8.1), the Lorentz force law Eq. (8.3) forms the foundation of electrodynamics, optics, and electric circuits. All classic electrodynamic phenomena can be described by these five equations.

$$\boldsymbol{F} = q\,(\boldsymbol{E} + \boldsymbol{v} \times \boldsymbol{B}) \qquad \text{(Lorentz force law)} \tag{8.3}$$

There are only slowly varying processes within the actuator compared to electromagnetic wave propagation. Furthermore, the conduction current $\boldsymbol{j}$ overbalances the displacement current $\frac{\partial \boldsymbol{D}}{\partial t}$ [120]. In addition, the capacitance of the system is very small, which means that the energy stored in the electric field of the coil is negligible against the magnetic field. Hence, the displacement current $\frac{\partial \boldsymbol{D}}{\partial t}$ in Ampere-Maxwell's law (Eq. (8.1c)) is set to zero [128]:

$$\operatorname{rot} \boldsymbol{H} = \boldsymbol{j}\,. \tag{8.4}$$

This assumption means not that all fields are regarded as static. Time dependence persists by Faraday's law of induction (Eq. (8.1a)) and current density $\boldsymbol{j} = \boldsymbol{j}(t)$ which is a function of time.
Integration of Eq. (8.1a) over area A yields the electromotive force or voltage of induction due to a changing magnetic induction:

$$U_{ind} = -\frac{\mathrm{d}}{\mathrm{d}\,t} \int \boldsymbol{B}\,\mathrm{d}A\,. \tag{8.5}$$

Magnetic Vector Potential

In many cases problems can be solved more easily if electrodynamic potentials are used instead of the field quantities. The magnetic induction $\boldsymbol{B}$ is a solenoidal or

divergence free vector field (Eq. (8.1d)) implying that the $\boldsymbol{B}$ field can be written as the curl of a vector potential $\boldsymbol{A}$ [13]:

$$\boldsymbol{B} = \operatorname{rot} \boldsymbol{A} \,. \tag{8.6}$$

As the curl of an arbitrary gradient field vanishes ($\operatorname{rot} \operatorname{grad} \Psi = 0 \,\forall \Psi = \Psi(\boldsymbol{r})$) [27], $\operatorname{grad} \Psi$ can be added to the vector potential $\boldsymbol{A}$ without having any effect on $\operatorname{rot} \boldsymbol{A}$. To ensure uniqueness, an additional gauge condition has to be stated. In this work the Coulomb gauge is used:

$$\operatorname{div} \boldsymbol{A} = 0 \,. \tag{8.7}$$

Combining the simplified Ampere-Maxwell law (Eq. (8.4)) with the definition of permeability (Eq. (8.2b)) and the Coulomb gauge (Eq. (8.7)) yields the Poisson equation:

$$\boldsymbol{\Delta} A = -\mu \boldsymbol{j} \,. \tag{8.8}$$

In cylindrical coordinates the Poisson equation for the vector potential $\boldsymbol{A} = (A_r \; A_\varphi \; A_z)$ is written as:

$$\Delta A_r - \frac{2}{r^2} \frac{\partial A_\varphi}{\partial \varphi} - \frac{A_r}{r^2} = -\mu j_r \tag{8.9a}$$

$$\Delta A_\varphi + \frac{2}{r^2} \frac{\partial A_r}{\partial \varphi} - \frac{A_\varphi}{r^2} = -\mu j_\varphi \tag{8.9b}$$

$$\Delta A_z = -\mu j_z \tag{8.9c}$$

with the Laplace operator in cylindrical coordinates:

$$\Delta = \frac{1}{r} \frac{\partial}{\partial r} r \frac{\partial}{\partial r} + \frac{1}{r^2} \frac{\partial^2}{\partial \varphi^2} + \frac{\partial^2}{\partial z^2} \,.$$

Boundary Conditions

At the interface between two media with different permeabilities μ_1 and μ_2, the fields $\boldsymbol{B}$ and $\boldsymbol{H}$ must fulfil additional restraints, called boundary conditions, which result from applying Stokes' and Gauss' theorem to the magnetic part of Maxwell's equations (Eq. (8.1d) and (8.4)) [56]:

$$\boldsymbol{B}_{n2} = \boldsymbol{B}_{n1} \tag{8.10}$$

$$(\boldsymbol{H}_2 - \boldsymbol{H}_1) \times \boldsymbol{n} = \boldsymbol{K} \,. \tag{8.11}$$

First, the magnetic induction B_n in normal direction $\boldsymbol{n}$ is continuous. Second, the tangential components of $\boldsymbol{H}$ are continuous across the interface for a current free surface. It otherwise changes discontinuously by an amount given by the surface current density $\boldsymbol{K}$ (current per length).

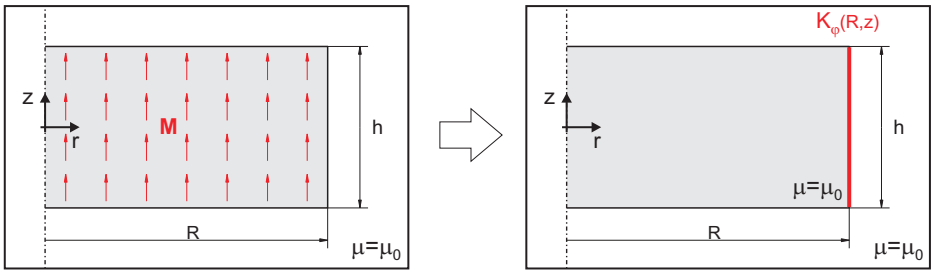


Figure 8.5: The homogeneous magnetisation of a cylindrical permanent magnet is equivalent to an azimuthal surface density

Permanent Magnetic Fields

Consider a homogeneous magnetic field from a cylindrical permanent magnet with radius R, height h, and magnetisation $\boldsymbol{M}_1$ in z direction (Figure 8.5).

According to Eq. (8.4), it is possible to introduce a magnetisation current $\boldsymbol{j}_M$ that is equivalent to the magnetisation field strength $\boldsymbol{M}$:

$$\operatorname{rot} \boldsymbol{M} = \boldsymbol{j}_M \,. \tag{8.12}$$

For rotational symmetry and magnetisation $\boldsymbol{M} = (0 \;\; 0 \;\; M_z)$ in z direction, the magnetisation current has only an azimuthal component $j_{M,\varphi}$:

$$\operatorname{rot} \boldsymbol{M} = \begin{pmatrix} 0 \\ -\frac{\partial M_z}{\partial r} \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ j_{M,\varphi} \\ 0 \end{pmatrix} \,. \tag{8.13}$$

The magnetisation is given by $M_z = M_1$ inside the cylinder and by $M_z = 0$ outside the cylinder. According to Eq. 8.10, this discontinuous shift is equal to the surface current density $K_{M,\varphi}$, which is proportional to a hypothetical current I:

$$K_{M,\varphi}(r,z) = \begin{cases} M_1 = \frac{I}{h} & \text{if } r = R \text{ and } |z| \leq \frac{h}{2} \\ 0 & \text{else} \,. \end{cases} \tag{8.14}$$

With the hypothetical current density, the electromagnet and permanent magnet can be described in the same way. Hence, the equations derived in this section are valid for both types of magnets.

Magnetic Force

Two different methods can be used to calculate the magnetic force: the Maxwell stress tensor and the Lorentz force law.

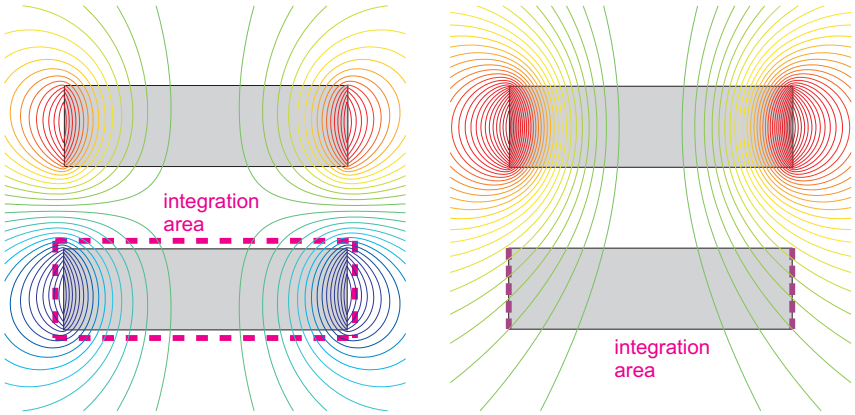


Figure 8.6: Magnetic induction and integration area of Maxwell stress tensor (left) and Lorentz force law (right)

Maxwell Stress Tensor

First, the common magnetic induction $\boldsymbol{B}$ of two magnets is calculated by superposition. Then, the magnetic force can be calculated by integrating the Maxwell stress tensor [54, 126] along a closed surface around one of the magnets (see left part of Figure 8.6). The magnetic force density $\boldsymbol{f}$ is given by Eq. (8.15):

$$\boldsymbol{f} = (\boldsymbol{j} + \operatorname{rot} \boldsymbol{M}) \times \boldsymbol{B} = \operatorname{div} \boldsymbol{T}_M \,. \tag{8.15}$$

The tensor

$$\boldsymbol{T}_M = \frac{1}{\mu_0} \boldsymbol{B} \otimes \boldsymbol{B}$$

is known as the magnetic part of the Maxwell stress tensor. Making use of Gauss's divergence theorem the volume integral (volume Ω with surface area $\partial\Omega$) is written as surface integral. The magnetic force $\boldsymbol{F}_M$ is given by:

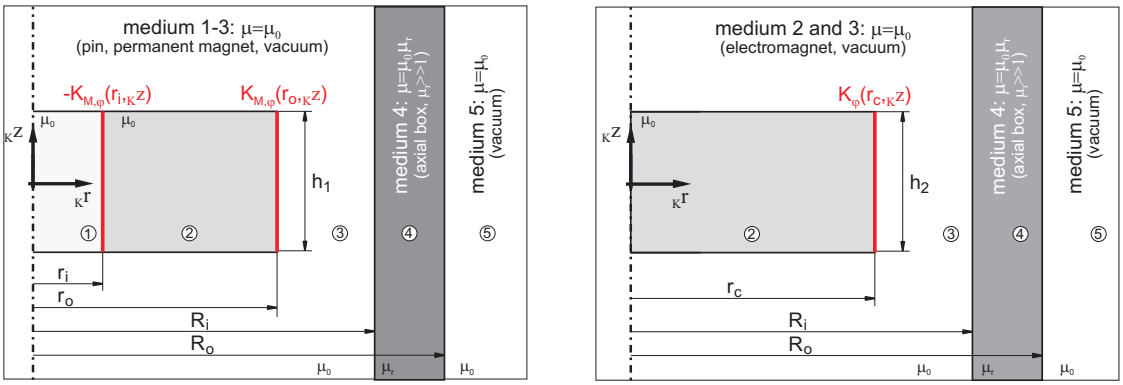
$$\boldsymbol{F}_M = \int\limits_{\Omega} \operatorname{div} \boldsymbol{T}_M \, \mathrm{d}V = \oint\limits_{\partial\Omega} \boldsymbol{n}\boldsymbol{T}_M \, \mathrm{d}A = \oint\limits_{\partial\Omega} \frac{1}{2\mu_0} \left(2(\boldsymbol{B}\boldsymbol{n})\boldsymbol{B} - B^2 \boldsymbol{n} \right) \mathrm{d}A \,. \tag{8.16}$$

Lorentz Force Law

If all electric currents—even the hypothetical ones, which are caused by the magnetisation of materials—are known, the magnetic force can also be calculated by the Lorentz force law (Eq. (8.17)) [127].

$$\boldsymbol{F}_M = \int \boldsymbol{K}_M \times \boldsymbol{B} \, \mathrm{d}A \tag{8.17}$$

$\boldsymbol{B}$ is the magnetic induction of the first magnet, whereas $\boldsymbol{K}_M$ is the surface current density of the second magnet. The magnetic force results from the Lorentz force on moving charges of the second magnet caused by the magnetic field of the first magnet. To get the force by integration (right part of Figure 8.6), only the magnetic induction $\boldsymbol{B}$ of the first magnet has to be evaluated on the lateral surface of the



(a) Permanent magnet (b) Electro magnet

Figure 8.7: Boundary conditions field equations

second magnet. As the vector potential of the first magnet has to be evaluated only on a few points, this method is numerically much cheaper than the Maxwell stress tensor method.

8.3.2 Magnetic Field of Actuator

The task is to calculate the vector potential of the magnetic fields for a given electric current under consideration of certain boundary and symmetry conditions. The boundary conditions for the permanent magnet of the actuator are given in Eq. (8.18) and illustrated in Figure 8.7. Analogue conditions are valid for the electromagnet.

$$
\begin{aligned}
H_{z,2} - H_{z,1} &= -K_{M,\varphi} && \text{boundary conditions at } r = r_i \\
B_{r,2} - B_{r,1} &= 0 \\
H_{z,3} - H_{z,2} &= K_{M,\varphi} && \text{boundary conditions at } r = r_{o,c} \\
B_{r,3} - B_{r,2} &= 0 \\
H_{z,4} - H_{z,3} &= 0 && \text{boundary conditions at } r = R_i \\
B_{r,4} - B_{r,3} &= 0 \\
H_{z,5} - H_{z,4} &= 0 && \text{boundary conditions at } r = R_o \\
B_{r,5} - B_{r,4} &= 0
\end{aligned}
\tag{8.18}
$$

Fields with Rotational Symmetry

The rotational symmetry of the actuator results in $\frac{\partial}{\partial \varphi} = 0$, $B_\varphi = 0$, $j_r = 0$, while $j_z = 0$. j_φ is the only component of the current density that does not vanish [115]. Consequently, the vector potential $\boldsymbol{A}$ only consists of the azimuthal component

$A_\varphi(r,z)$. The magnetic induction $\boldsymbol{B}$ (Eq. (8.6)) reduces to:

$$\boldsymbol{B} = \left(-\frac{\partial A_\varphi}{\partial z} \quad 0 \quad \frac{1}{r}\frac{\partial}{\partial r}(rA_\varphi)\right)^T . \tag{8.19}$$

The rotational symmetry reduces the Poisson equation (8.9) to:

$$\frac{1}{r}\frac{\partial}{\partial r}r\frac{\partial}{\partial r}A_\varphi(r,z) + \frac{\partial^2 A_\varphi(r,z)}{\partial z^2} - \frac{A_\varphi(r,z)}{r^2} = \mu j_\varphi(r,z) . \tag{8.20}$$

The space with electric current is split into a current free space and the necessary boundary conditions containing the surface current density. Hence, the homogeneous part of Eq. (8.20) is solved by separating the variables:

$$A_\varphi(r,z) = R(r)Z(r) . \tag{8.21}$$

The partial differential equation is split into two ordinary differential equations by this ansatz. With separation of eigenvalue k, the ordinary differential equation for $Z(z)$ is written as:

$$\frac{\partial^2 Z(z)}{\partial z^2} = -k^2 Z(z), \qquad k \in \mathbb{R}_0^+ \tag{8.22}$$

with solution

$$Z(z) = A_1(k)\cos(kz) + A_2(k)\sin(kz) . \tag{8.23}$$

The ordinary differential equation for $R(r)$ becomes the modified Bessel's differential equation of first order:

$$\frac{1}{r}\frac{\partial}{\partial r}r\frac{\partial}{\partial r}R(r) - \left(k^2 + \frac{1}{r^2}\right)R(r) = 0 . \tag{8.24}$$

The solution of the modified Bessel's differential equation can be expressed by a linear combination of the modified Bessel function of first $I_1(kr)$ and second kind $K_1(kr)$ of first order [27]:

$$R(r) = \tilde{C}_1(k)I_1(kr) + \tilde{C}_2(k)K_1(kr) . \tag{8.25}$$

The solution for the azimuthal component of the vector potential A_φ gets

$$A_\varphi(r,z) = \int_0^\infty \left(A_1(k)\cos(kz) + A_2(k)\sin(kz)\right)\left(\tilde{C}_1(k)I_1(kr) + \tilde{C}_2(k)K_1(kr)\right)\,\mathrm{d}k . \tag{8.26}$$

Because of the symmetry equation for the surface current density:

$$K_{M,\varphi}(z) = K_{M,\varphi}(-z) , \tag{8.27}$$

and the rotational symmetry of the actuator, terms containing $\sin(kz)$ are set to zero. Furthermore, the limits of the modified Bessel functions are divergent:

$$\lim_{x\to\infty} I_1(x) = \infty \tag{8.28}$$

$$\lim_{x\to 0} K_1(x) = \infty\,. \tag{8.29}$$

Hence, the integration constants $\tilde{C}_i$ coming along with $I_1(kr)$ must vanish in the most outer region and with $K_1(kr)$ in the most inner region. To solve the boundary value problem (Eqs. (8.18) and (8.20)), the ansatz below is used.

$$\begin{aligned}
A_\varphi^{(1)}(r,z) &= \int_0^\infty C_1(k) I_1(kr)\cos(kz)\,\mathrm{d}k && r \le r_i \\
A_\varphi^{(2)}(r,z) &= \int_0^\infty \left(C_2(k) I_1(kr) + C_3(k) K_1(kr)\right)\cos(kz)\,\mathrm{d}k && r_i < r \le r_o \\
A_\varphi^{(3)}(r,z) &= \int_0^\infty \left(C_4(k) I_1(kr) + C_5(k) K_1(kr)\right)\cos(kz)\,\mathrm{d}k && r_o < r \le R_1 \\
A_\varphi^{(4)}(r,z) &= \int_0^\infty \left(C_6(k) I_1(kr) + C_7(k) K_1(kr)\right)\cos(kz)\,\mathrm{d}k && R_i < r \le R_o \\
A_\varphi^{(5)}(r,z) &= \int_0^\infty C_8(k) K_1(kr)\cos(kz)\,\mathrm{d}k && r > R_o
\end{aligned} \tag{8.30}$$

This ansatz contains eight unknown integration constants $C_i(k)$ $i = 1...8$, which have to be calculated by the corresponding boundary conditions at the three lateral surfaces. The boundary conditions of Eq. (8.18) yield:

$$\begin{aligned}
H_{z,2} - H_{z,1} &= \frac{1}{\mu_0}\left(\frac{1}{r}\frac{\partial}{\partial r} r A_\varphi^{(2)}(r,z) - \frac{1}{r}\frac{\partial}{\partial r} r A_\varphi^{(1)}(r,z)\right)\bigg|_{r=r_i} = -K_{M,\varphi} \\
B_{r,2} - B_{r,1} &= -\frac{\partial A_\varphi^{(2)}(r_i,z)}{\partial z} + \frac{\partial A_\varphi^{(1)}(r_i,z)}{\partial z} = 0 \\
H_{z,3} - H_{z,2} &= \frac{1}{\mu_0}\left(\frac{1}{r}\frac{\partial}{\partial r} r A_\varphi^{(3)}(r,z) - \frac{1}{r}\frac{\partial}{\partial r} r A_\varphi^{(2)}(r,z)\right)\bigg|_{r=r_o} = -K_{M,\varphi} \\
B_{r,3} - B_{r,2} &= -\frac{\partial A_\varphi^{(3)}(r_o,z)}{\partial z} + \frac{\partial A_\varphi^{(2)}(r_o,z)}{\partial z} = 0 \\
H_{z,4} - H_{z,3} &= \frac{1}{\mu_0\mu_r}\,\frac{1}{r}\frac{\partial}{\partial r} r A_\varphi^{(4)}(r,z)\bigg|_{r=R_i} - \frac{1}{\mu_0}\frac{1}{r}\frac{\partial}{\partial r} r A_\varphi^{(3)}(r,z)\bigg|_{r=R_i} = 0 \\
B_{r,4} - B_{r,3} &= -\frac{\partial A_\varphi^{(4)}(R_i,z)}{\partial z} + \frac{\partial A_\varphi^{(3)}(R_i,z)}{\partial z} = 0 \\
H_{z,5} - H_{z,4} &= \frac{1}{\mu_0\mu_r}\,\frac{1}{r}\frac{\partial}{\partial r} r A_\varphi^{(5)}(r,z)\bigg|_{r=R_o} - \frac{1}{\mu_0}\frac{1}{r}\frac{\partial}{\partial r} r A_\varphi^{(4)}(r,z)\bigg|_{r=R_o} = 0
\end{aligned}$$

$$B_{r,5} - B_{r,4} \quad = - \frac{\partial\, A_{\varphi}^{(5)}(R_o,z)}{\partial\, z} + \frac{\partial\, A_{\varphi}^{(4)}(R_o,z)}{\partial\, z} = 0\,. \tag{8.31}$$

To evaluate Eq. (8.31), the recurrence relations and properties of Bessel's functions stated below are used [2].

$$\frac{\mathrm{d}}{\mathrm{d}\,x}\left(xI_1(x)\right) = xI_0(x) \tag{8.32}$$

$$\frac{\mathrm{d}}{\mathrm{d}\,x}\left(xK_1(x)\right) = -xK_0(x) \tag{8.33}$$

$$I_0(x)K_1(x) + K_0(x)I_1(x) = \frac{1}{x} \tag{8.34}$$

Furthermore, the surface current density at $r = r_o$ is analysed by Fourier expansion:

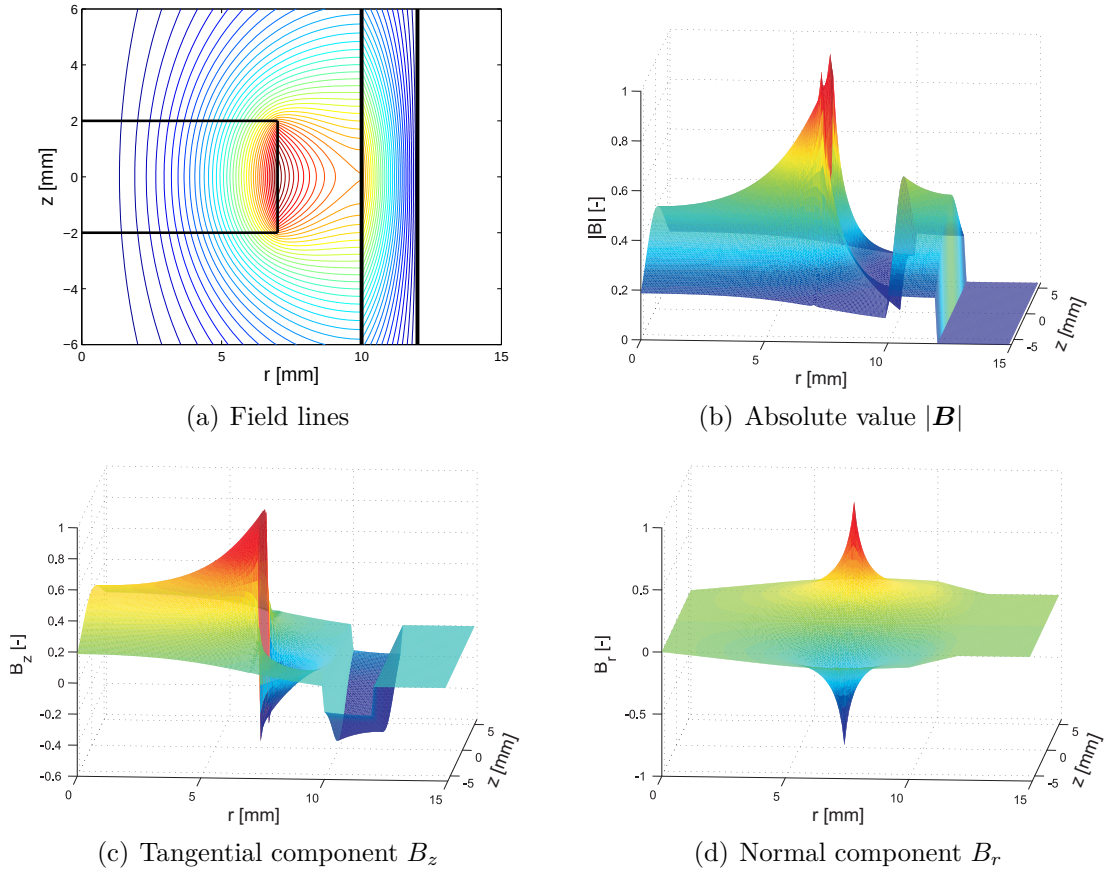
$$\begin{aligned} K_{M,\varphi}(z) = M_1\,\mathrm{rect}\left(\frac{z}{h_1}\right) &= M_1\left(\int_0^{\infty} 2\frac{\sin(\frac{h_1}{2}k)}{\pi k}\cos(kz)\mathrm{d}k - \frac{h_1}{2\pi}\right) \\ &\approx M_1 \int_0^{\infty} 2\frac{\sin(\frac{h_1}{2}k)}{\pi k}\cos(kz)\mathrm{d}k \end{aligned} \tag{8.35}$$

$$\mathrm{rect}(x) = \begin{cases} 0 & \text{if } |x| > \frac{1}{2} \\ \frac{1}{2} & \text{if } |x| = \frac{1}{2} \\ 1 & \text{if } |x| < \frac{1}{2} \end{cases}. \tag{8.36}$$

Substituting the vector potential $A_{\varphi}^{(i)}$ in Eq. (8.31) with Eq. (8.30) under consideration of Eqs. (8.32) to (8.34), and equating coefficients with the Fourier synthesised surface current density (Eq. (8.35)), the boundary conditions result in a system of linear equations for $C_i(k)$:

$$\begin{aligned} &C_2(k)I_0(kr_i) - C_3(k)K_0(kr_i) - C_1(k)I_0(kr_i) = -2\mu_0 M_1 \frac{\sin(\frac{h_1}{2}k)}{\pi k^2} \\ &C_2(k)I_1(kr_i) + C_3(k)K_1(kr_i) - C_1(k)I_1(kr_i) = 0 \\ &C_4(k)I_0(kr_o) - C_5(k)K_0(kr_o) - C_2(k)I_0(kr_o) + C_3(k)K_0(kr_o) = 2\mu_0 M_1 \frac{\sin(\frac{h_1}{2}k)}{\pi k^2} \\ &C_4(k)I_1(kr_o) + C_5(k)K_1(kr_o) - C_2(k)I_1(kr_o) - C_3(k)K_1(kr_o) = 0 \\ &C_6(k)I_0(kR_i) - C_7(k)K_0(kR_i) - \mu_r C_4(k)I_0(kR_i) + \mu_r C_5(k)K_0(kR_i) = 0 \\ &C_6(k)I_1(kR_i) + C_7(k)K_1(kR_i) - C_4(k)I_1(kR_i) - C_5(k)K_1(kR_i) = 0 \\ &-\mu_r C_8(k)K_0(kR_o) - C_6(k)I_0(kR_o) + C_7(k)K_0(kR_o) = 0 \\ &C_8(k)K_1(kR_o) - C_6(k)I_1(kR_o) - C_7(k)K_1(kR_o) = 0\,. \end{aligned} \tag{8.37}$$

The solution for the unknown integration constants $C_i(k)$ is unique. The equations above describe the magnetic induction of the permanent magnet with boundary



(a) Field lines

(b) Absolute value $|\boldsymbol{B}|$

(c) Tangential component B_z

(d) Normal component B_r

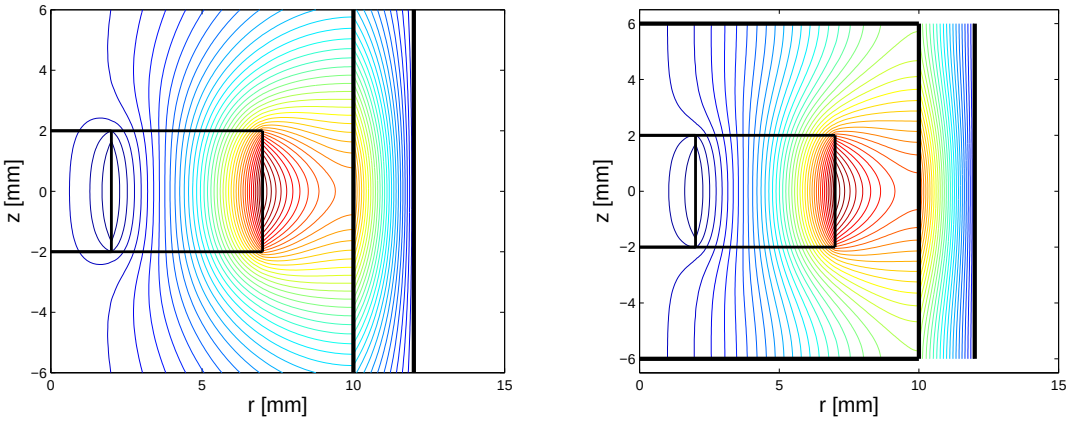
Figure 8.8: Magnetic field of the electromagnet in a radial box

conditions according to Figure 8.7(a).
The magnetic induction of the electromagnet can be obtained analogously or by calculating the limit of $\boldsymbol{B}$ for $r_i \to 0$. The normalised magnetic induction $\boldsymbol{B}$ of the electromagnet is shown in Figure 8.8.

In Lang [127] a similar system is solved with an alternative method using complete elliptic integrals. Using Fourier transformation, the equivalence of both methods can be proved.

To model an axial ferromagnetic box, the reflecting method [180] is used. A reflection of the magnetic field without an axial box is calculated and positioned with the same distance to the axial box, but opposite sign. The superposition of the original field with the reflection field yields the magnetic induction of the magnet with an axial box as an additional boundary condition. In Figure 8.9 the magnetic induction of the permanent magnet is plotted without an axial box (left) and with upper and lower axial boxes (right).

The field lines with constant magnetic induction $\boldsymbol{B}$ are bent towards the radial and axial boxes. Due to the high permeability μ of the boxes, the field lines hit them almost perpendicularly. At the outer radius of the radial box the magnetic induction almost vanishes.



(a) Magnetic field only with radial boundary conditions

(b) Magnetic field with radial and axial boundary conditions

Figure 8.9: Field lines of the permanent magnet in a radial box

8.3.3 Electrodynamics

In the previous section the magnetic induction $\boldsymbol{B}$ of the coil has been derived as a function of the current I. To calculate the electrodynamic current, the electric circuit shown in Figure 8.10 is used. It comprises the interior resistance R_i, the coil resistance R_L, the inductance L of the coil, and the shut down resistance R_e.

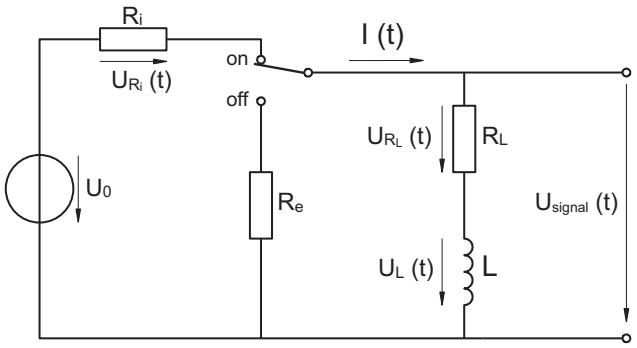


Figure 8.10: Electric circuit

Kirchhoff's circuit laws [36]

$$
\begin{aligned}
&U_{signal} = U_{R_L} + U_L \\
&I = I_{R_i} = I_{R_L} = I_L \\
&U = U_{R_i} + U_{R_L} + U_L \qquad \text{with } U = \begin{cases} U_0 & \text{switch } \textit{on} \\ 0 & \text{switch } \textit{off} \end{cases}
\end{aligned} \tag{8.38}
$$

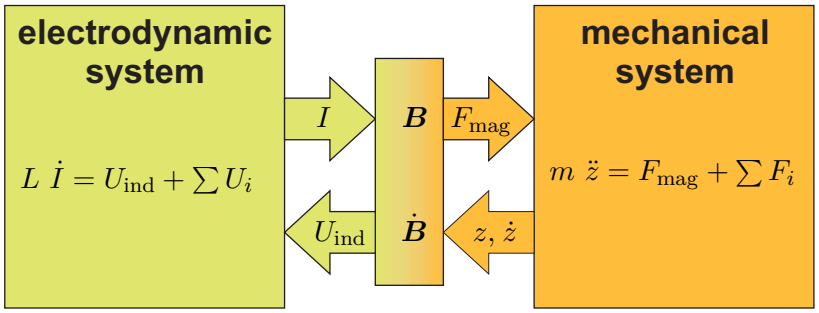


Figure 8.11: Coupling of mechanical and electrodynamic systems

in combination with the relations for the electronic components, coil voltage, and resistance voltage

$$U_R = RI\,, \tag{8.39}$$

$$U_L = L\frac{\mathrm{d}\,I}{\mathrm{d}\,t} \tag{8.40}$$

are used to describe the electrodynamics of the system in terms of an ordinary differential equation for the current I.

8.3.4 Coupling Mechanical and Electrodynamic Systems

The permanent magnet is attached to the pin. Hence, the motion of the pin causes a change in the magnetic flux fed by the permanent magnet at the coil. As seen in Figure 8.4(a), this motion results in an inductive voltage. To take this effect into account, Eq. (8.40) is expanded starting with Faraday's law (Eq. (8.5)):

$$\begin{aligned} U_{\text{ind}} = -\frac{\mathrm{d}}{\mathrm{d}\,t}\int \boldsymbol{B}\ \mathrm{d}A &= -\int \frac{\partial\,\boldsymbol{B}_{\text{coil}}}{\partial\,t}\,\mathrm{d}A - \int \frac{\partial\,\boldsymbol{B}_{\text{ext.}}}{\partial\,z}\,\dot{z}\,\mathrm{d}A \\ &= \quad L\frac{\mathrm{d}\,I}{\mathrm{d}\,t} \quad - \quad \frac{\partial\,\Phi_{\text{ext.}}}{\partial\,z}\,\dot{z}\,. \end{aligned} \tag{8.41}$$

The magnetic induction $\boldsymbol{B}_{\text{coil}}$ is proportional to the current I. Hence, the first term in Eq. (8.41) describes the self-inductance of the coil. The second term represents the induction due to a changing external field $\boldsymbol{B}_{\text{ext.}}$. For the actuator, $\boldsymbol{B}_{\text{ext.}}$ is fed by the permanent magnet and changes with its motion.

The coupling of both physical domains, mechanics and electrodynamics, via the magnetic induction is shown in Figure 8.11. The effect of coupling both differential equations on the electric system is that the pin motion generates an additional induction voltage. The consequence for the mechanical system is that the pin motion is dampened by induced eddy currents.

The time profile of the voltage U_{signal} at the clamps is plotted in Figure 8.12. The effect of induction due to the pin motion on U_{signal} is small during extending the pin.

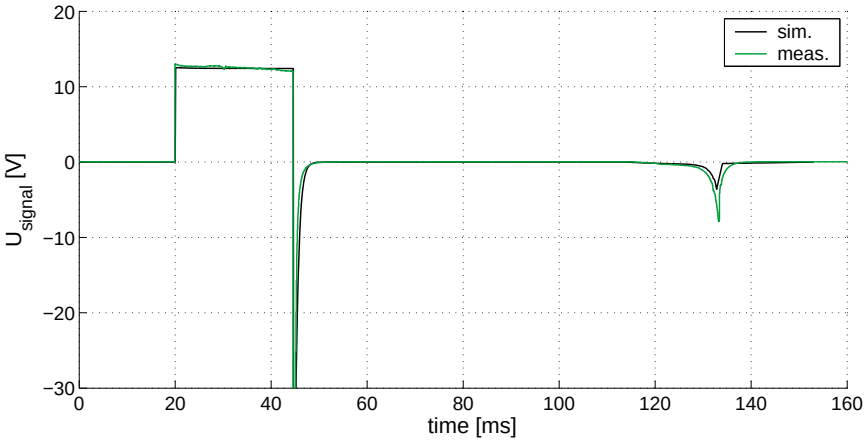


Figure 8.12: Induced voltage due to shut down and pin motion: simulation and measurement

Table 8.2: Static magnetic forces

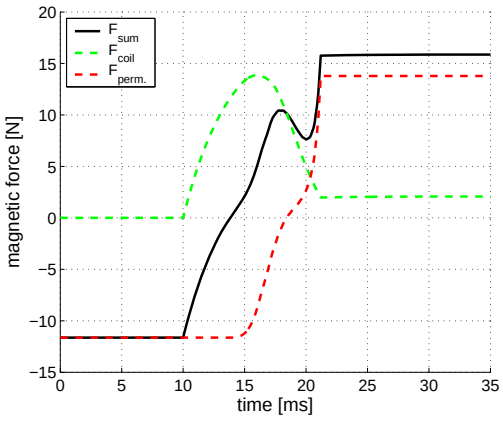
pin position		0 V	8 V	12 V
z= 0 mm (pin *in*)	simulation	-11.6 N	0.5 N	6.5 N
	measurement	-11.6 N	2.4 N	7.2 N
z= 4 mm (pin *out*)	simulation	13.8 N	15.2 N	15.9 N
	measurement	13.8 N	15.1 N	15.4 N

At $t \approx 130$ms the induction peak due to retracting the pin is quite visible as the circuit is open at this time.

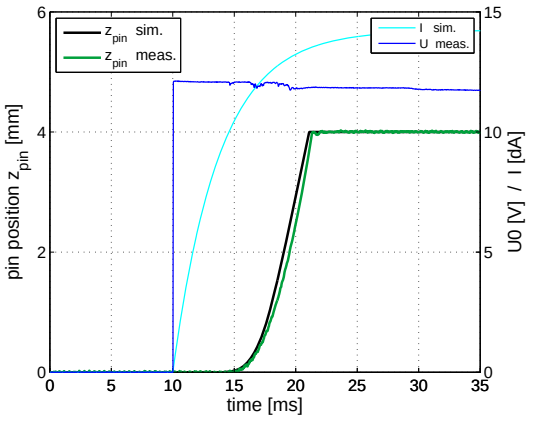
8.4 Results

To verify the derived simulation model and identify unknown parameters (e.g. permeability of the boxes, magnetisation of the permanent magnet and internal resistances), the simulation results are compared to experimental data.
To identify and adjust the simulation parameters, the static common magnetic forces at both end stops are measured first. The experimental data and simulation results for different, constantly connected load voltages are presented in Table 8.2. Except for the values at $z = 0$ and 8V input all deviations are smaller than 1N.

Next, voltage steps of 8V and 12V are applied and the position of the pin and voltage U at the clamps are measured (for test rig setup compare Figure 8.15(b)). The right parts of Figures 8.13 and Figure 8.14 present the measured pin position z_{pin} and voltage U, as well as the simulated pin position and current I. There is very good correlation between the pin motion in simulation and experiment for 12V,. For 8V, the reaction or dead time is about 2ms longer in the simulation than experimental data, whereas the flight time corresponds well. This is partly due to the fact that

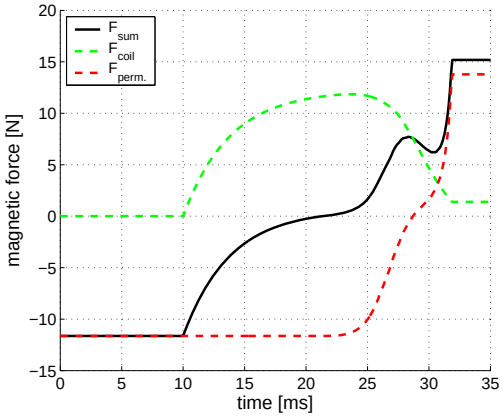


(a) Magnetic force components due to permanent magnet and solenoid

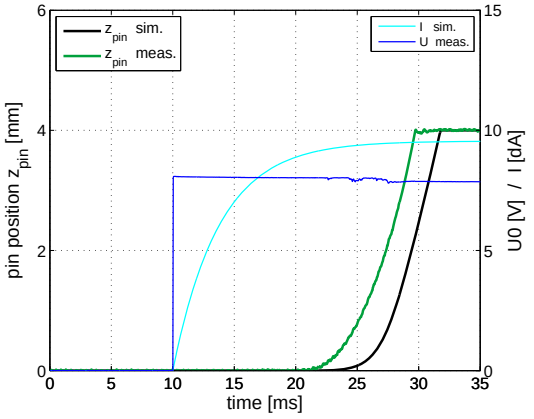


(b) Pin position in simulation and experiment

Figure 8.13: Supply voltage: 12V



(a) Magnetic force components due to permanent magnet and solenoid



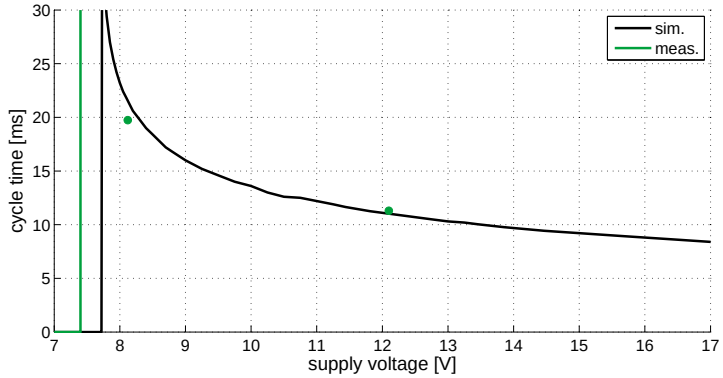
(b) Pin position in simulation and experiment

Figure 8.14: Supply voltage: 8V

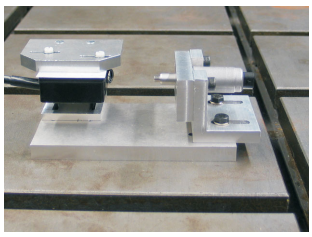
the common magnetic force of the simulation model is slightly too small for $z_{\text{pin}} = 0$ (Table 8.2) and partly due to the nature of the magnetic force, which rises very slow at 8V (Figure 8.14(a)).

In Figure 8.13(a) and Figure 8.14(a) the magnetic forces are plotted during the extending process of the magnetic actuator. The magnetic force $F_{\text{perm.}}$ due to the permanent magnet is a function of the distance between the two magnets. At both end stops the permanent magnet sticks to the axial ferromagnetic boxes. F_{coil} is the contribution of the solenoid. It is a function of the distance between both magnets and the electric current passing through the coil. At the beginning F_{coil} is rising, because of the increasing current. At the end, the current is almost constant and F_{coil} is decreasing with the distance becoming higher. The actual active force on the moving parts F_{sum} is the sum of $F_{\text{perm.}}$ and F_{coil}.

At the beginning of the changeover at $t = 10\text{ms}$, an 8V and 12V step are applied



(a) Time period for extending the pin as a function of the applied voltage



(b) Test rig actuator

Figure 8.15: Shifting times

respectively. The pin does not react directly on the step. This so-called dead time is caused by the iron core that is magnetised by the permanent magnet. The electromagnetic force F_{coil} of the coil has to overcome the magnetic force of the permanent magnet $F_{\text{perm.}}$ before the moving assembly starts to accelerate. Comparing the zero-crossing of the active force F_{sum} for 8V and 12V, the slope for 8V is very small. As a consequence, the dead time is very sensitive to force errors for low voltages.

In Figure 8.15(a) the flight time from the inner to outer end stop is plotted as a function of the applied voltage. For 12V the error between measured and simulated switching time is approximately 2.5 %. For lower voltages the error rises due to reasons mentioned previously. The critical voltage at which the pin does not move is reproduced accurately in the simulation to 0.3V.

9 Valve Springs

9.1 General Introduction to Valve Springs

The valve spring is one of the most important parts of the valve train regarding dynamic behaviour, friction, and operational reliability. A failure or fracture of this part may result in a total loss of the engine. Considering the dynamics, the valve spring is the component with the lowest natural frequency in the motor and therefore plays a crucial role in the overall dynamics of the valve train.

9.1.1 Characteristics of Valve Springs

The purpose of the valve spring is to store energy and maintain the contact between components in the valve train so that the valve motion will follow the cam profile. The valve spring keeps the valve tightly closed against its seat until the cam opens the valve and contracts the spring. At the maximum valve lift, the kinetic energy of the valve is stored in the valve spring and is used to retract the valve into its seat. The spring force must be high enough to close the valve reliably and prevent the valve from bouncing off the seating due to surge modes after it has been closed. Conversely, the spring forces affect the friction level in the engine and therefore fuel consumption. For this reason the spring force should be kept as low as possible.

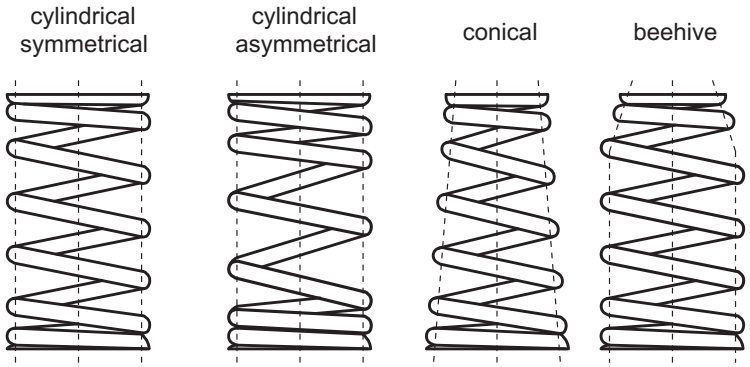


Figure 9.1: Winding shapes of valve springs

Different types of valve springs are used to generate the spring force. Helical coil springs with different winding shapes are usually used [14]. Other types of springs used especially in the early days of engine design are hairpin, leaf, and torsion bar springs [18,110]. Due to the limited space in the cylinder head of modern multi-valve

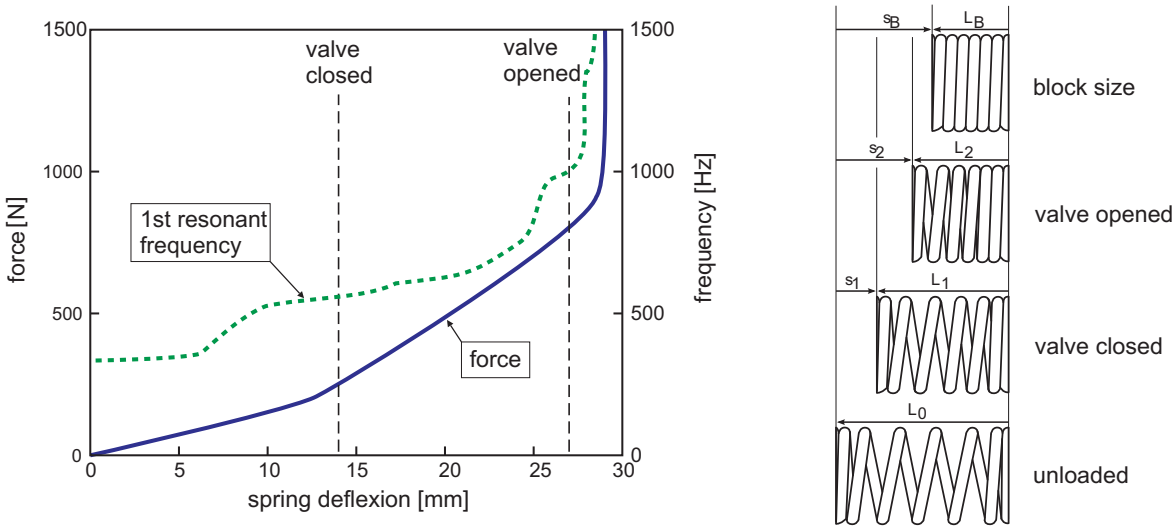


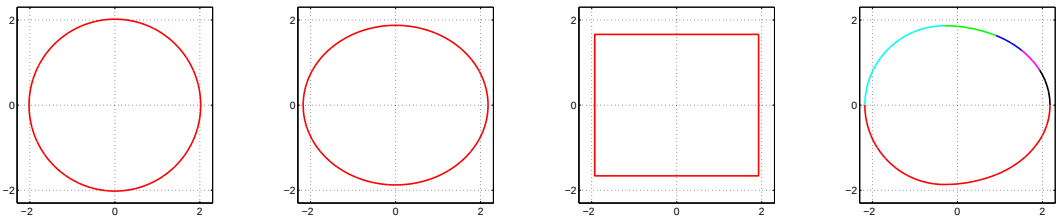
Figure 9.2: Non-linear behaviour of valve springs

engines, these types of springs are not commonly used. Another interesting approach are pneumatic valve springs, for example, in high-speed Formula One racing engines. Pneumatic valve springs enable high rotational speeds of up to 20,000 rpm due to their progressive characteristic and the reduced number of moving parts [182].

Some typical winding shapes of coil springs are shown in Figure 9.1. Nearly without exception valve springs of current engines have a progressive behaviour. This behaviour is mainly caused by a non-constant pitch between adjacent coils. The coils with a lower pitch come into contact earlier than coils with a higher pitch. The active part of the spring is reduced resulting in increasing total spring stiffness with increasing compression of the spring. The characteristic becomes progressive and resonant frequencies rise (Figure 9.2).

An operating point near the natural frequency can result in excessive vibrations and forces. This phenomenon is called surging and has a negative effect on the durability of springs and causes valves to bounce off their seating after they have closed. Moreover, if the forces are great enough, adjacent coils may clash. Springs with asymmetric shapes are less prone to surging because of lower moving masses and geometrical reasons [101].
For further reduction of the moving mass, conical valve springs are used. Conical springs have advantages that the moving mass of the spring itself is reduced, smaller spring fixings and collars are possible, and the solid length is marginally shorter. A drawback is the less progressive characteristics compared to cylindrical springs. Beehive springs are a combination of cylindrical and conical springs. The moving mass is reduced by the conical part and the progressive characteristic is provided by the cylindrical part.

The cross section of a valve spring wire is not restricted to a circular shape. Some typical cross sections are shown in Figure 9.3. Elliptical cross sections allow a reduction of the overall height. Multi-arc profiles (MA) are generated from several circular segments with differing radii and centre points. Compared to elliptical cross



(a) circular (b) elliptical (c) rectangular (d) multi arc

Figure 9.3: Wire cross section shapes

sections, MA profiles display better stress distribution and allow better utilisation of the material. For example, Muhr [149] uses MA profiles in combination with beehive winding shapes to reduce the spring forces in a valve train by 13%.

9.1.2 Impact on Fuel Consumption

Although there is a public demand for cars with lower specific fuel consumption, additional costs are only accepted on a small scale. Manufacturing costs of valve springs are comparatively low and therefore valve springs provide an economical starting point for the reduction of fuel consumption.
In the lower rpm range in combustion engines approximately 30% of the friction losses are caused by the valve train. Under full load conditions, this contribution reduces to 8% and in return the friction at the pistons and crankshaft becomes dominant [184]. Since vehicles are mostly operated under partial load, fuel consumption can be reduced significantly by reducing the spring forces in valve trains. In Speckens [184] an example is given for improvement of the fatigue limit by 7% that results in a 400 rpm higher limit for the valve train. If the engine is operated at the same rpm range, this improvement can be used to reduce the spring force to result in a consumption advantage of 0.3%.

All measures to reduce the spring force contribute to an optimisation of engine efficiency. In general, these measures aim to reduce moving masses or improve material utilisation. Some possibilities are listed below:

- beehive valve spring
- elliptical cross section
- multi-arc cross sections
- high tensile steel
- lightweight construction material (e.g. titan [150])

Besides friction, valve springs also affect fuel consumption by their influence on the valve lift. To achieve an optimal gas exchange and minimise flow losses, the periods for opening and closing the valves, especially for opening the exhaust valve, should be minimal. Regarding this, a rectangular valve lift would be desirable. As a square wave contains infinite odd harmonics, the bandwidth of the excitation is very large.

Hence, the dynamic behaviour and natural vibrations of valve springs become more and more important if the valve trajectory draws near to a rectangular shape.

9.1.3 Literature Survey

As a very important and widespread machine part, coil springs have been the focus of research since the beginning of the 20th century. First relevant works concerning stiffness and stress distribution in spring wires were done by GÖHNER and go back to the 1930s [78–80]. A summary of the works concerning stress distribution of non-circular cross sections is given by NIEPAGE [154] and an overall summary of metal springs is given by MEISSNER and SCHORCHT [140].

The dynamics of coil springs were first studied using a one dimensional wave equation (e.g. HUSSMANN [8] or PISCHINGER [166]). A brief overview is provided by SPECKENS [184]. Under the assumption of small rotations, WITTRICK [203] derived six coupled partial differential equations to describe the dynamics of coil springs. MOTTERSHEAD [146] presented a finite element based on the partial differential equations and derived an extension for large deflections [147]. PHILLIPS and COSTELLO [162] presented a solution of the linearised equations. NAGAYA et al. [151] analysed the free vibration of coil springs with arbitrary shape.
Using static or natural frequency analyses, all of the above-mentioned investigators showed that the set of six partial differential equations, which are also the topic of this work, provide an accurate way of modelling coil springs without coil contacts. Results from LEE [130] indicate that by modelling the spring as a curved beam and by considering motion in all directions, higher frequencies are more accurately modelled.

To get time-resolved results for arbitrary excitations, there are two possibilities: transient FEM analyses or considering the spring as a flexible body within a multi-body system. In the works of HÄUFLE [94] and SÄGER [174], detailed 3D FEM models of coil springs are used. As this approach requires a lot of computational time, it is not practical for investigating springs in valve train multi-body simulations. Mainly three dynamic spring-models are used within the context of multi-body simulations: modal model, multi-mass model and multi-beam model.

Modal Model

The motion of a flexible body can be described by superimposing its different eigenmodes. The eigenmodes are weighted with time-dependent functions. In most cases, it is sufficient to consider the first few eigenmodes because the most energy is contained in these modes. The main advantage of this spring model is that the number of degrees of freedom is small and therefore the computational time is relatively short. A modal-reduced model was used by PHILIPS et. al. [163] in valve train simulations. The serious disadvantage, however, is that no contact modelling is possible.

For that reason this model does not have enough potential for detailed simulations; for example it is not possible to map progressive characteristics.

Multi-Mass Model

The multi-mass model is most commonly used in valve train simulations. The valve spring is modelled by a series of masses, coupled with spring-damper-elements. More precisely, each coil is split into 4 to 8 masses. The stiffness of the spring can be calculated from a static experiment. The individual stiffness of each connecting spring can be calculated from the overall stiffness.
Advantages of the multi-mass model are that progressive behaviour can be modelled and coil interactions are possible. On the other hand, the model has two meaningful drawbacks: the amount of degrees of freedom is high and the equations of motion will mostly become numerically stiff, potentially resulting in a higher computational effort.
Bakaj [12] used a multi-mass model without contacts between the coils. A nonlinear multi-mass model with progressive characteristic was presented by Ross [172] and Speckens [184]. Non-smooth contacts between coils were added by Engelhardt [50]. Multi-mass models are also available in commercial multi-body simulation environments (e.g. FEV Virtual Engine (former MSC ADAMS / Engine) [155]) and are used for industrial applications [52].
An extended multi-mass model was presented by Wittkopp [202]. He considered the coil spring as an unwound torsion rod, which is discretised using rotational force elements. Lateral motion is possible with this parametrisation as the rotation of the cross section and translation in spring direction are uncoupled.

Multi-Beam Model

Straight or curved beams are used to approximate the spring wire in the multi-beam model. If straight beams are applied, a large amount of beams is needed to approximate the curvature of the spring wire. Tichanek et al. [189] used 24 elements per coil for a valve spring model within a valve train simulation. This results in a significantly higher computational effort compared to the multi-mass model.
Dreyer [43] developed a valve spring model based on a finite element for a curved beam with 12 degrees of freedom. Experimental analyses were performed by fixing strain gauges to a valve spring to assure the results of the simulation, [71].

9.1.4 Objectives and Structure

For modelling valve springs, the substantial difference between the dynamic response of the spring and the static response has to be taken into account. The internal dynamics of the spring, such as moving masses of the coils and contacts between the

windings, must be considered. Consequently, sophisticated valve spring models considering the internal dynamics are needed for the design and optimization of modern valve trains. This is important as in the future components such as valve springs will also need to be fully optimised in order to reduce fuel consumption.
In this thesis a new valve spring model for efficient multi-body simulations of valve trains is presented. The approach is to approximate the spring wire as a curved beam and derive the equations starting with the basic equations of continuum mechanics. The resulting system of six coupled partial differential equations is reduced to one hyperbolic partial differential equation to utilise the model for multi-body simulations. The interactions between windings are included using the Augmented Lagrangian method and non-smooth contact mechanics.
Compared to a multi-mass-model approach, the presented method allows the degrees of freedom and the stiffness of the resulting differential equations to be reduced. Therefore, the computational effort can be reduced along with slightly increased accuracy. The model is validated in the frequency and time domain with experimental data using a test rig.

9.2 Continuum Mechanics

In this section the theoretical background and methods to derive the governing equations for a coil spring are provided. For a more detailed presentation see [9,16,104]. First, the fundamental concepts and principles of continuum mechanics are presented. Continuum mechanics study the response of materials to different loading conditions. Starting with the kinematics of a continuum, strain as a measure to describe deformations is introduced. Using material-dependent constitutive equations, the strain is related to stress, resultant forces, and momentums. Applying balance laws and equilibrium equations such as the principles of linear and angular momentum yields the equation of motion as partial differential equations (PDE). To solve such equations numerically, the PDEs must be discretised. For this purpose the Finite Element Method (FEM) is used to yield ordinary differential equations that can be integrated easily in a common multi-body system approach.

9.2.1 Kinematics and Strain

A particle or material point within the context of continuum mechanics means an infinite volume of material inside a body or object.
Let us suppose that an undeformed body at certain time $t = t_0$ occupies a certain region Ω_0 of physical space. The position of a particle in the undeformed configuration can be described by its position vector $\boldsymbol{a}$. As the position in the undeformed state is unique for each particle, $\boldsymbol{a}$ is often used to label or reference a certain material point and is often known as the material coordinate.
The new shape of the deformed body at time t is denoted by Ω_t. The vector $\boldsymbol{x} = \boldsymbol{x}(\boldsymbol{a},t)$ references the position of a single particle $\boldsymbol{a}$ in the deformed body, while

$\boldsymbol{x}(\boldsymbol{a},t)$ describes the path line or trajectory of the particle and is also known as the spatial coordinate or deformation. The displacement $\boldsymbol{u}$ and the deformation gradient tensor $\boldsymbol{F}$ are defined as:

$$\boldsymbol{u}(\boldsymbol{a},t) = \boldsymbol{x}(\boldsymbol{a},t) - \boldsymbol{a} \tag{9.1}$$

$$\boldsymbol{F} = \nabla_a \boldsymbol{x} = \frac{\partial \boldsymbol{x}}{\partial \boldsymbol{a}}. \tag{9.2}$$

One of the basic tasks of continuum mechanics is to describe the strain within material objects. As the deformation gradient involves rigid body motion, $\boldsymbol{F}$ is not suitable for this purpose. By elimination of rotations and rigid body motion, the Green-Lagrangian or Eulerian-Almansi strain tensors are defined (see Figure 9.4).

Lagrangian and Eulerian Description

When a continuum is in motion, quantities associated with specific particles change with time. These changes can be described by two different perspectives.

Within the Lagrangian description, which is normally used in solid mechanics, the position and physical properties of the particles are described in terms of the material coordinate $\boldsymbol{a}$ and time. An observer is focused on the trajectory of a single particle. Physical properties are expressed as functions of $\boldsymbol{a}$ and time t. The material or substantial derivative of any property f is simply the partial derivative with respect to time (Eq. (9.3)) and $\boldsymbol{a}$ is held constant.

$$\frac{\mathrm{D}f}{\mathrm{D}t} = \dot{f} = \frac{\partial f(\boldsymbol{a},t)}{\partial t} \tag{9.3}$$

The Eulerian or spatial description focuses on the current configuration, giving attention to what is occurring at a fixed point in space. This approach is conveniently applied in fluid dynamics as flow lines are of more interest than trajectories of all particles. Physical properties are expressed as a function of $\boldsymbol{x}$ and time t. To obtain the material or substantial derivative of any property f, the chain rule has to be used. This yields the local and convective rate of change:

$$\frac{\mathrm{D}f(\boldsymbol{x},t)}{\mathrm{D}t} = \frac{\partial f}{\partial t} + \boldsymbol{v} \cdot \nabla_x f \qquad \text{with } \boldsymbol{v} = \frac{\partial \boldsymbol{x}}{\partial t}\,. \tag{9.4}$$

9.2.2 Kinetics and Equations of Motion

The subjects of this section are kinetic quantities such as forces, momentums, and stresses. Stress is a result of external loads applied to the surface or volume of a deformable continuum. The correlation between deformation (in terms of stress) and loads (in terms of strain) is given by constitutive equations.

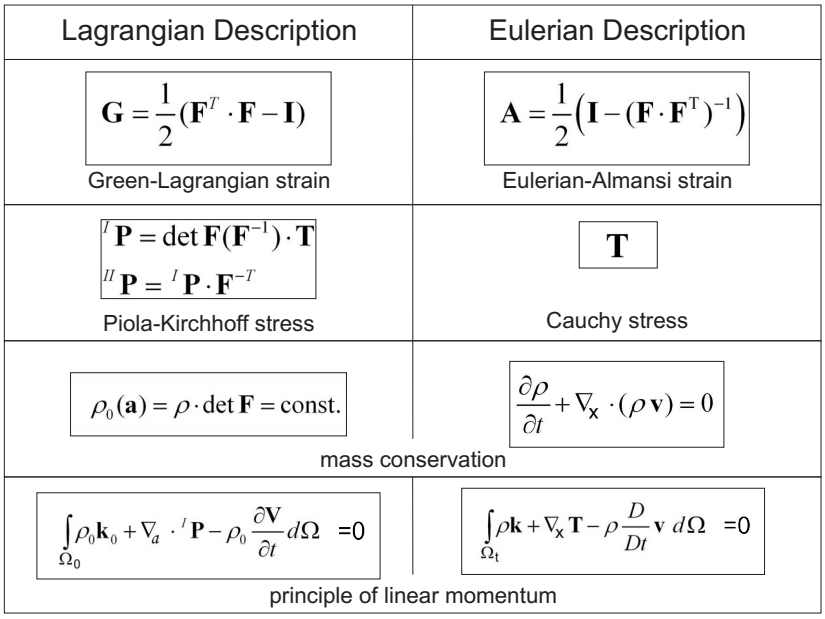


Figure 9.4: Lagrangian and Eulerian description: Overview

With the mass density ρ, the body force density $\boldsymbol{k}$, and the surface force vector per unit area $\boldsymbol{t}$, body and surface forces are defined by Eq. (9.5):

$$\begin{aligned} \mathrm{d}\boldsymbol{f}_1 &= \rho \cdot \boldsymbol{k} \cdot \mathrm{d}V \qquad \text{volume force} \\ \mathrm{d}\boldsymbol{f}_2 &= \boldsymbol{t} \cdot \mathrm{d}A \qquad\quad \text{surface force}\,. \end{aligned} \tag{9.5}$$

According to Cauchy, the stress at any point P and any tangent planes containing P is completely defined by nine components of a second-order tensor known as stress tensor. For the Eulerian description the stress components are related to the area in the deformed configuration. The corresponding stress tensor $\boldsymbol{T}$ is called the Cauchy stress tensor. It is used for stress analyses with small deformations. For large deformations other measures such as the first and second Piola-Kirchhoff or Kirchhoff stress tensors are required, which are related to the area of the reference configuration.

To relate forces and stresses to the motion of the continuous body the equations of motion must be derived. The basic postulate is that each particle of the continuum must satisfy Newton's law of motion. Starting with the principle of linear momentum for one particle and integrating it over the body volume Ω with corresponding surface $\partial\Omega$ and normal $\boldsymbol{n}$ yields:

$$\frac{\mathrm{d}}{\mathrm{d}\,t} \int\limits_{\Omega_t} \rho \boldsymbol{v} \mathrm{d}V = \int\limits_{\Omega_t} \rho \boldsymbol{k} \mathrm{d}\Omega + \int\limits_{\partial\Omega_t} \boldsymbol{n} \cdot \boldsymbol{T} \mathrm{d}A\,. \tag{9.6}$$

Figure 9.4 gives an overview of the continuum mechanical quantities of this section.

9.2.3 Linear Theory

For small displacement gradients $\nabla \boldsymbol{u} << 1$, which do not necessarily mean that the displacements $\boldsymbol{u}$ are small, geometric linear theory may be applied and the difference between Lagrangian and Eulerian description vanishes.

$$\boldsymbol{\varepsilon} = \frac{1}{2}\left(\nabla_{\boldsymbol{a}} \boldsymbol{u}^T + \nabla_{\boldsymbol{a}} \boldsymbol{u}\right) \tag{9.7}$$

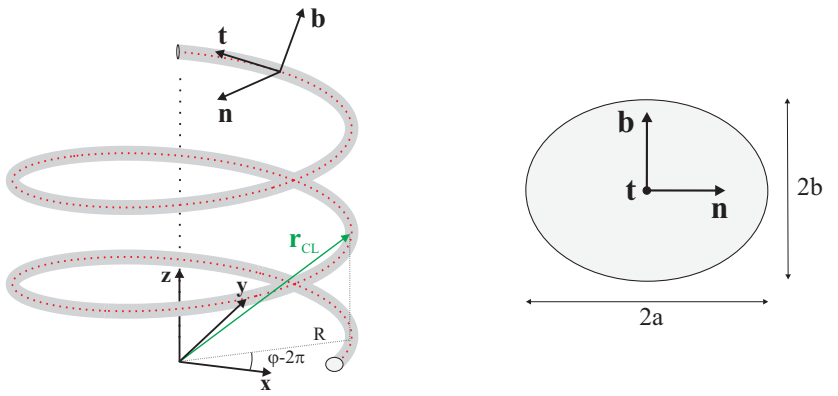
Geometric linear problems are generally solved with the linear strain tensor $\boldsymbol{\varepsilon}$ as defined in Eq. (9.7) and the Cauchy stress tensor $\boldsymbol{T} = \boldsymbol{\sigma}$. The constitutive law used in this work is Hooke's generalised law (Eq. (9.8)) with Young's modulus E and lateral contraction coefficient μ.

$$\begin{aligned}
\sigma_{tt} &= \frac{E}{1+\mu}\left(\varepsilon_{tt} + \frac{\mu}{1-2\mu}(\varepsilon_{tt} + \varepsilon_{nn} + \varepsilon_{bb})\right) \\
\sigma_{nn} &= \frac{E}{1+\mu}\left(\varepsilon_{nn} + \frac{\mu}{1-2\mu}(\varepsilon_{tt} + \varepsilon_{nn} + \varepsilon_{bb})\right) \\
\sigma_{bb} &= \frac{E}{1+\mu}\left(\varepsilon_{bb} + \frac{\mu}{1-2\mu}(\varepsilon_{tt} + \varepsilon_{nn} + \varepsilon_{bb})\right) \\
\tau_{tn} &= \frac{E}{(1+\mu)}\varepsilon_{tn} \\
\tau_{tb} &= \frac{E}{(1+\mu)}\varepsilon_{tb} \\
\tau_{nb} &= \frac{E}{(1+\mu)}\varepsilon_{nb}
\end{aligned} \tag{9.8}$$

Constitutive equations for different materials can be found in literature for general non-linear cases [9, 11, 194].

9.3 Governing Equations of the Coil Spring

In this section the partial differential equations of motion for a mass carrying coil spring is derived. In a first step the coil spring is approximated as a curved beam. In accordance with regular beam theory, the overall displacement is composed of displacement of the centre line, rotation of cross sections, and warping. Based on this, the strain measurements of an arbitrary point of the wire can be determined. The stress in the spring coil is calculated as a function of the strain using Hooke's linear constitutive law. Six coupled partial differential equations are derived by applying the principles of linear and angular momentum. To make the model more comfortable to use within multi-body simulation, these six equations are reduced to a one-dimensional wave equation at the end of this section.



(a) Centre line of the coil spring with Cartesian and Frenet-Serret frame

(b) Elliptical cross section with major and minor semi-axis a,b

Figure 9.5: Geometry of coil spring and used frames

9.3.1 Kinematic Equations

From a geometrical point of view, a coil spring can be described by a cross section extruded along a smooth space curve (=centre line) that is wound about a central axis. The cross section is assumed to be elliptical and constant. As a consequence, the geometry of the coil spring can be defined by a helical centre line $\boldsymbol{r}_{CL}$ of the winding and the semi-axis a, b of the elliptical cross section (see Figure 9.5).

The centre line can be parametrised in cylindrical coordinates (R,φ,z) by the angle φ:

$$\boldsymbol{r}_{CL}(\varphi) = \begin{pmatrix} R(\varphi)\cos\varphi \\ R(\varphi)\sin\varphi \\ \frac{h(\varphi)}{2\pi}\varphi \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} . \tag{9.9}$$

R is the radius of the coil curve after projection onto the xy-plane and h is a quantity describing the pitch of the curve. For constant R and h, h denotes the distance between two coils. For valve springs the pitch and radius are not constant, but a function of φ.

Frenet-Serret Frame

Instead of a regular Cartesian coordinate system, tangent, normal and binormal are used as an orthonormal basis for a new reference frame at each point on the curve. Such a frame is called a Frenet-Serret frame. To calculate the basis of the

Frenet-Serret frame the curve is parametrised by its arc length s instead of φ.

$$s(\varphi) = \int_{\varphi_0}^{\varphi} \left\| \frac{\mathrm{d}\,\boldsymbol{r}_{CL}(\xi)}{\mathrm{d}\,\xi} \right\| \mathrm{d}\xi, \qquad \boldsymbol{t} = \frac{\mathrm{d}\,\boldsymbol{r}_{CL}}{\mathrm{d}\,s}, \quad \boldsymbol{n} = \frac{\mathrm{d}\,\boldsymbol{t}}{\mathrm{d}\,s} \frac{1}{\left\| \frac{\mathrm{d}\,\boldsymbol{t}}{\mathrm{d}\,s} \right\|}, \quad \boldsymbol{b} = \boldsymbol{t} \times \boldsymbol{n} \tag{9.10}$$

For coil springs it can be assumed that $\boldsymbol{r}_{CL}(\varphi)$ is a unique parametrisation and $\frac{\mathrm{d}\,\boldsymbol{r}_{CL}(\varphi)}{\mathrm{d}\,\varphi} \neq 0$. Therefore, Eq. (9.10) can be inverted yielding φ as a function of s.

Derivatives in the Frenet-Serret Frame

The curvature κ and torsion τ are often used to characterize a continuous differentiable space curve. The curvature at a special point of a given curve is the reciprocal of the radius of the osculating circle at that point. The torsion of a curve measures how sharply it is twisting. The relation between the derivatives of the Frenet-Serret basis and the quantities torsion and curvature are given by the Frenet-Serret theorem [27]. In finding the derivative of a vector expressed in the Frenet-Serret frame, it must be considered that the Frenet-Serret frame changes with s. An arbitrary vector $\boldsymbol{a}$ with components a_t, a_n, and a_b in the Frenet-Serret frame is written as:

$$\boldsymbol{a} = a_t(s) \cdot \boldsymbol{t}(s) + a_n(s) \cdot \boldsymbol{n}(s) + a_b(s) \cdot \boldsymbol{b}(s)\,. \tag{9.11}$$

Making use of the Frenet-Serret theorem, the derivative of $\boldsymbol{a}$ can be calculated. For a compact presentation the vector-matrix notation as shown in the next equation is often very useful:

$$\frac{\mathrm{d}}{\mathrm{d}\,s} \begin{pmatrix} a_t \\ a_n \\ a_b \end{pmatrix} = \begin{pmatrix} \frac{\mathrm{d}\,a_t}{\mathrm{d}\,s} - \kappa a_n \\ \frac{\mathrm{d}\,a_n}{\mathrm{d}\,s} + \kappa a_t - \tau a_b \\ \frac{\mathrm{d}\,a_b}{\mathrm{d}\,s} + \tau a_n \end{pmatrix}. \tag{9.12}$$

Displacement and Strain

A cross section of the spring wire in the undeformed state is characterised by its corresponding point P on the centre line $\boldsymbol{r}_{0P}(s)$. As illustrated in Figure 9.6, Q is an arbitrary point in the cross section and can be referenced with the coordinates s, η, and ζ (Eq. (9.13)):

$$\boldsymbol{r}_{0Q}(s,\eta,\zeta) = \boldsymbol{r}_{0P}(s) + \eta\boldsymbol{n}(s) + \zeta\boldsymbol{b}(s)\,. \tag{9.13}$$

It is assumed that the total displacement $\boldsymbol{u}$ can be composed of three components:

- translation of the centre line
- small rotation of the cross section
- warping of the cross section

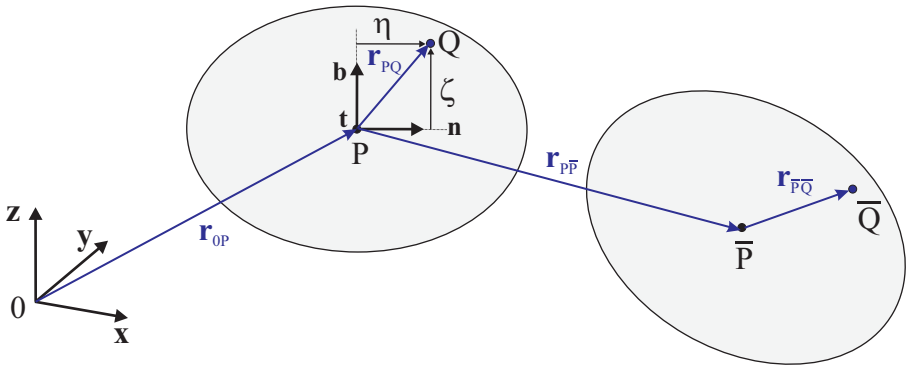


Figure 9.6: Deformed state: Reference to an arbitrary point $\bar{Q}$ after deformation due to translation and rotation

The translation and rotation can be interpreted as rigid body motion of independent cross sections, whereas warping is an out-of-plane displacement of the cross section. Translation and rotation are illustrated in Figure 9.6. Quantities in the deformed state are labelled with a bar ($^{-}$).

Translation

The translation of any point of the cross section is equal to the translation of the centre line. The deformation can be expressed in the Frenet-Serret frame:

$$\boldsymbol{r}_{P\bar{P}} = u_t(s)\boldsymbol{t}(s) + u_n(s)\boldsymbol{n}(s) + u_b(s)\boldsymbol{b}(s)\,. \tag{9.14}$$

As the centre line is parametrised by only one parameter s, the coordinates u_t, u_n, u_b, and the translation $\boldsymbol{r}_{P\bar{P}}$ itself are also functions only of the arc length s.

Rotation

As rotation is assumed to be small, the order of the elemental rotations can be neglected. For instance, Cardan angles can be used to describe the rotation by its rotation matrix $\boldsymbol{A}$ which is composed of three elemental rotations:

$$\boldsymbol{A} = \boldsymbol{A}_b(\vartheta_b)\boldsymbol{A}_n(\vartheta_n)\boldsymbol{A}_t(\vartheta_t)\,. \tag{9.15}$$

For small rotations Eq. (9.15) is linearised and the rotation of an arbitrary point in the cross section can be calculated:

$$\boldsymbol{r}_{\bar{P}\bar{Q}} = \boldsymbol{A} \cdot \boldsymbol{r}_{PQ} = (\vartheta_n\zeta - \vartheta_b\eta)\boldsymbol{t} + (\eta - \vartheta_t\zeta)\boldsymbol{n} + (\vartheta_t\eta + \zeta)\boldsymbol{b}\,. \tag{9.16}$$

Warping

Torsion applied to a non-circular cross section causes the cross section to not only rotate, but also warp. Warping is an out-of-plane displacement in $\boldsymbol{t}$-direction of the cross section as illustrated in Figure 9.7.

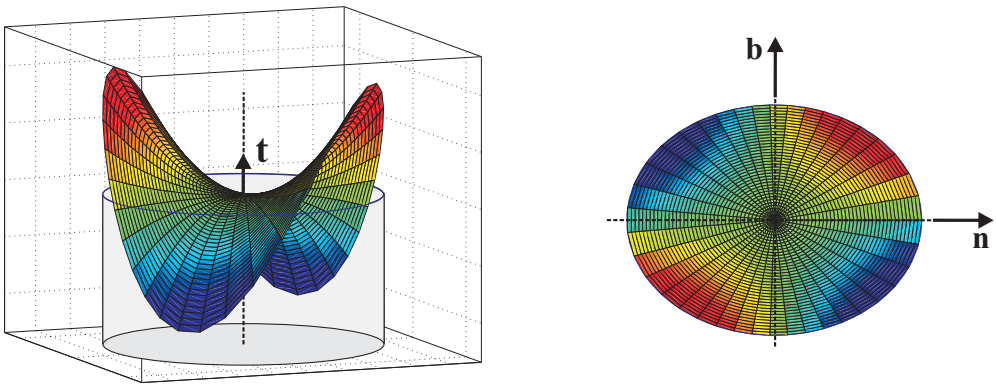


Figure 9.7: Deformed state: Deformation due to torsional warping of an elliptical cross section

Overestimation of the torsional rigidity of non-circular cross sections results when warping is neglected. The deformation due to warping as a result of twisted cross sections is given by Eq. (9.17):

$$\boldsymbol{u}_{\text{warping}} = \psi(\eta,\zeta)\alpha(s)\boldsymbol{t} \tag{9.17}$$

where $\psi(\eta,\zeta)$ is the warping function and $\alpha(s)$ is a quantity describing the twist of the cross section. Assuming that the curvature of the wire has no effect on the warping behaviour, the warping function Eq. (9.18) from regular theory for a straight beam is valid:

$$\psi\,(\eta,\zeta) = \frac{b^2 - a^2}{a^2 + b^2}\eta\zeta\,. \tag{9.18}$$

Displacement

According to Figure 9.6 an arbitrary point in the deformed state is referenced by:

$$\boldsymbol{r}_{0\bar{Q}} = \boldsymbol{r}_{0P} + \boldsymbol{r}_{P\bar{P}} + \boldsymbol{r}_{\bar{P}\bar{Q}} + \boldsymbol{u}_{\text{warping}}\,. \tag{9.19}$$

The difference of the deformed and undeformed state results in the displacement $\boldsymbol{u}$ in the Frenet-Serret frame:

$$\boldsymbol{u}(s,\eta,\zeta) = \begin{pmatrix} u_t(s) + \vartheta_n(s)\zeta - \vartheta_b(s)\eta + \psi(\eta,\zeta)\alpha(s)) \\ u_n(s) - \vartheta_t(s)\zeta \\ u_b(s) + \vartheta_t(s)\eta \end{pmatrix}. \tag{9.20}$$

As the deformations are assumed to be small, a linear strain measurement based on geometric linear theory is feasible and the linear Cauchy strain as defined in Eq. (9.7) can be calculated:

$$\varepsilon_{tt} = \frac{\mathrm{d}\,u_t}{\mathrm{d}\,s} - \kappa u_n + \zeta\left(\frac{\mathrm{d}\,\vartheta_n}{\mathrm{d}\,s} + \kappa\vartheta_t - \tau\vartheta_b\right) - \eta\left(\frac{\mathrm{d}\,\vartheta_b}{\mathrm{d}\,s} + \tau\vartheta_n\right) + \psi\frac{\mathrm{d}\,\alpha}{\mathrm{d}\,s}$$

$$
\begin{aligned}
\varepsilon_{tn} &= \frac{1}{2}\left(\frac{\mathrm{d}\,u_n}{\mathrm{d}\,s} + \kappa u_t - \tau u_b - \zeta\left(\frac{\mathrm{d}\,\vartheta_t}{\mathrm{d}\,s} - \kappa\vartheta_n\right) - \vartheta_b + \frac{\mathrm{d}\,\psi}{\mathrm{d}\,\eta}\alpha\right)\\
\varepsilon_{tb} &= \frac{1}{2}\left(\frac{\mathrm{d}\,u_b}{\mathrm{d}\,s} + \tau u_n + \eta\left(\frac{\mathrm{d}\,\vartheta_t}{\mathrm{d}\,s} - \kappa\vartheta_n\right) + \vartheta_n + \frac{\mathrm{d}\,\psi}{\mathrm{d}\,\zeta}\alpha\right)\\
\varepsilon_{nn} &= 0\\
\varepsilon_{bb} &= 0\\
\varepsilon_{nb} &= 0\,.
\end{aligned}
\tag{9.21}
$$

By comparing the derived strain tensor with a straight beam [98],

$$
\alpha(s) = \frac{\mathrm{d}\,\vartheta_t}{\mathrm{d}\,s} - \kappa\vartheta_n \tag{9.22}
$$

is chosen as the quantity determining the twist of the cross section.

9.3.2 Constitutive Equations

The Cauchy strain tensor Eq. (9.21) and Hooke's generalised law are used to derive the forces acting on the cross sections.
As deformation of the wire has been parametrised without considering lateral contraction, applying Hooke's generalised law as stated in Eq. (9.8) would result in false stresses in $\boldsymbol{n}$ and $\boldsymbol{b}$ direction, making the wire too stiff as $\varepsilon_{nn} = 0$ and $\varepsilon_{bb} = 0$ must also be fulfilled. Therefore, Hooke's generalised law is reduced to Eq. (9.23) using $\varepsilon_{nn} = \varepsilon_{bb} = -\mu\varepsilon_{tt}$:

$$
\begin{aligned}
\sigma_{tt} &= E\varepsilon_{tt}\\
\tau_{tn} &= \frac{E}{(1+\mu)}\varepsilon_{tn} = 2G\varepsilon_{tn}\\
\tau_{tb} &= \frac{E}{(1+\mu)}\varepsilon_{tb} = 2G\varepsilon_{tb}
\end{aligned}
\tag{9.23}
$$

with the shear modulus

$$
G = \frac{E}{2(1+\mu)}\,. \tag{9.24}
$$

The resultant forces F_t, F_n, F_b, and moments M_t, M_n, and M_b at the cross section are obtained by integrating the stresses over the cross section A. For an elliptical cross section the constitutive equations are given in Eq. (9.25):

$$
\begin{aligned}
F_t &= EA\left(\frac{\mathrm{d}\,u_t}{\mathrm{d}\,s} - \kappa u_n\right)\\
F_n &= GA\left(\frac{\mathrm{d}\,u_n}{\mathrm{d}\,s} + \kappa u_t - \tau u_b - \vartheta_b\right)
\end{aligned}
$$

$$F_b = GA\left(\frac{\mathrm{d}\,u_b}{\mathrm{d}\,s} + \tau u_n + \vartheta_n\right)$$

$$\begin{aligned} M_t &= GI_t\left(\frac{\mathrm{d}\,\vartheta_t}{\mathrm{d}\,s} - \kappa\vartheta_n\right) \\ M_n &= EI_\eta\left(\frac{\mathrm{d}\,\vartheta_n}{\mathrm{d}\,s} + \kappa\vartheta_t - \tau\vartheta_b\right) \\ M_b &= EI_\varsigma\left(\frac{\mathrm{d}\,\vartheta_b}{\mathrm{d}\,s} + \tau\vartheta_n\right) \end{aligned} \tag{9.25}$$

with the torsional second moment of area

$$I_t = \frac{a^3b^3}{a^2+b^2}\pi\,. \tag{9.26}$$

9.3.3 Equations of Motion

The constitutive equations Eq. (9.25), which are describing the relation between displacement and resultant forces, are used to derive the equations of motion of an individual cross section. The following assumptions are made:

- curvature is negligible for the leverage of the momentums
- distributed loads are constant within the infinitesimal part
- motion due to warping has no dynamic effects
- rotations are small

With the distributed loads $\boldsymbol{f}$ and $\boldsymbol{d}$, Eq. (9.6) is evaluated yielding Eq. (9.27)

$$EA\left(\frac{\partial^2 u_t}{\partial s^2} - \frac{\partial\,\kappa u_n}{\partial s}\right) - \kappa GA\left(\frac{\partial\,u_n}{\partial s} + \kappa u_t - \tau u_b - \vartheta_b\right) + f_t = \rho A\frac{\partial^2 u_t}{\partial t^2}$$

$$\begin{aligned} &GA\left(\frac{\partial^2 u_n}{\partial s^2} + \frac{\partial\,\kappa u_t}{\partial s} - \frac{\partial\,\tau u_b}{\partial s} - \frac{\partial\,\vartheta_b}{\partial s}\right) + \kappa EA\left(\frac{\partial\,u_t}{\partial s} - \kappa u_n\right) - \\ &\tau GA\left(\frac{\partial\,u_b}{\partial s} + \tau u_n + \vartheta_n\right) + f_n = \rho A\frac{\partial^2 u_n}{\partial t^2} \end{aligned}$$

$$\begin{aligned} &GA\left(\frac{\partial^2 u_b}{\partial s^2} + \frac{\partial\,\tau u_n}{\partial s} + \frac{\partial\,\vartheta_n}{\partial s}\right) + \tau GA\left(\frac{\partial\,u_n}{\partial s} + \kappa u_t\right) - \\ &\tau GA\left(\tau u_b + \vartheta_b\right) + f_b = \rho A\frac{\partial^2 u_b}{\partial t^2} \end{aligned}$$

$$
\begin{aligned}
&GI_t\left(\frac{\partial^2\vartheta_t}{\partial s^2}-\frac{\partial\kappa\vartheta_n}{\partial s}\right)-\kappa EI_\eta\left(\frac{\partial\vartheta_n}{\partial s}+\kappa\vartheta_t-\tau\vartheta_b\right)+d_t && =\rho AR^2_{g_{tt}}\frac{\partial^2\vartheta_t}{\partial t^2}\\
&EI_\eta\left(\frac{\partial^2\vartheta_n}{\partial s^2}+\frac{\partial\kappa\vartheta_t}{\partial s}-\frac{\partial\tau\vartheta_b}{\partial s}\right)+\kappa GJ\left(\frac{\partial\vartheta_t}{\partial s}-\kappa\vartheta_n\right) && -\\
&\tau EI_\zeta\left(\frac{\partial\vartheta_b}{\partial s}+\tau\vartheta_n\right)-GA\left(\frac{\partial u_b}{\partial s}+\tau u_n+\vartheta_n\right)+d_n && =\rho AR^2_{g_{nn}}\frac{\partial^2\vartheta_n}{\partial t^2}\\
&EI_\zeta\left(\frac{\partial^2\vartheta_b}{\partial s^2}+\frac{\partial\tau\vartheta_n}{\partial s}\right)+\tau EI_\eta\left(\frac{\partial\vartheta_n}{\partial s}+\kappa\vartheta_t-\tau\vartheta_b\right)+\\
&+GA\left(\frac{\partial u_n}{\partial s}+\kappa u_t-\tau u_b-\vartheta_b\right)+d_b && =\rho AR^2_{g_{bb}}\frac{\partial^2\vartheta_b}{\partial t^2}.
\end{aligned}
\tag{9.27}
$$

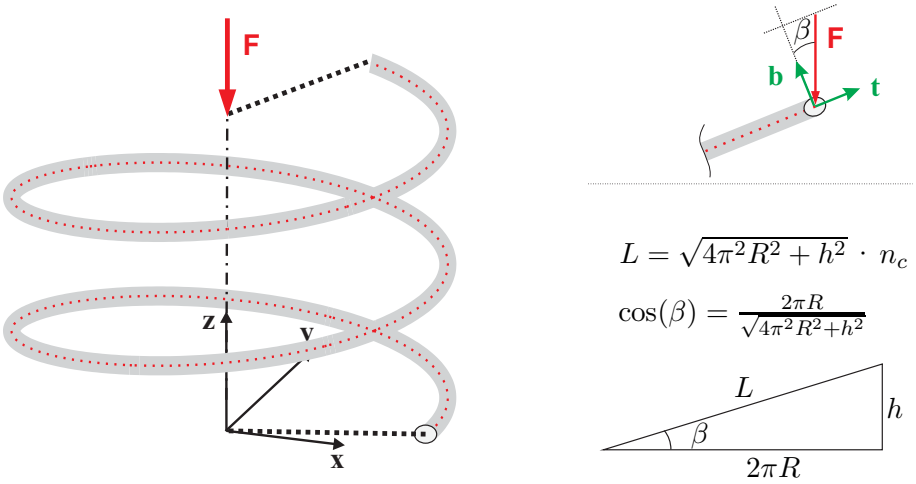
The equations of motion for the curved beam are derived as six coupled partial differential equations. The equations are valid for constant cross sections. If the cross section varies along s, the derivatives of A, E, G, I_η, I_ζ, and I_t with respect to s have to be considered.
Similar equations have already been specified by WITTRICK [203]. In contrast to this work, WITTRICK derives the forces and momentums directly using an infinite element and neglects warping.

9.3.4 Reducing the Partial Differential Equations

The aim of this work is to derive a set of ordinary differential equations that can be incorporated into a multi-body simulation framework. As the main focus of this work are valve springs that are mainly under axial load, further restrictions for movement of the cross sections are admissible to reduce the degrees of freedom and save computational time. This is done by examining a coil spring under axial load. By calculating the spring stiffness, torsion is identified as the main influencing quantity and the equations are reduced to one torsional degree of freedom.
A coil spring under axial load is depicted in Figure 9.8. The force $\boldsymbol{F}$ is parallel to the centre line and applied to the spring via virtual rigid extensions. The lower end is clamped. The spring has n_c number of active coils, radius R, constant pitch h, and an elliptical cross section with semi-axis a and b.

Decomposition of the external axial force $\boldsymbol{F}$ yields the internal forces and momentums at the cross section:

$$
\begin{aligned}
\boldsymbol{F}_{ax,t} &= -\sin\beta F & \boldsymbol{M}_{ax,t} &= -R\sin\beta F\\
\boldsymbol{F}_{ax,n} &= 0 & \boldsymbol{M}_{ax,n} &= 0\\
\boldsymbol{F}_{ax,b} &= -\cos\beta F & \boldsymbol{M}_{ax,b} &= -R\cos\beta F\,.
\end{aligned}
\tag{9.28}
$$



(a) Coil spring with axial load (b) Pitch of coil spring

Figure 9.8: Coil spring with constant pitch and axial load along the centre line

Elastic Energy and Approximation of Spring Stiffness

Starting with elastic energy U, the spring stiffness can be derived with Castigliano's theorem [31]:

$$U = \int_A \int_0^L \frac{1}{2}\boldsymbol{\sigma} : \boldsymbol{\varepsilon} \; \mathrm{d}s \, \mathrm{d}A \quad = \int_A \int_0^L \frac{E}{2}\varepsilon_{tt}^2 + \frac{E}{1+\mu}\varepsilon_{tn}^2 + \frac{E}{1+\mu}\varepsilon_{tb}^2 \; \mathrm{d}s \, \mathrm{d}A \,. \tag{9.29}$$

For the static case, a uniform torsion of spring wire can be assumed. Next, the strains are substituted with Eq. (9.21) and integrated. With Eq. (9.25), Eq. (9.28), and the abbreviation $\hat{\mu} = 2\cos\beta(1+\mu)$, the elastic energy can be written as:

$$\begin{aligned} U_{ax} = & \quad \frac{1}{2}\int_0^L \left(\frac{F_{ax,t}^2}{EA} + \frac{F_{ax,b}^2}{GA} + \frac{M_{ax,t}^2}{GI_t} + \frac{M_{ax,b}^2}{EI_\zeta} \right) ds \\ = & \frac{2\pi R^3 n_c}{2GI_t} F^2 \left(\frac{\sin^2\beta I_t}{\hat{\mu} A R^2} + \frac{\cos\beta I_t}{AR^2} + \cos\beta + \frac{\sin^2\beta I_t}{\hat{\mu} I_\zeta} \right). \end{aligned} \tag{9.30}$$

According to Castigliano's theorem the displacement w in the direction of the force F at the free end is given by the partial derivative:

$$w = \frac{\partial U_{ax}}{\partial F} = \frac{2\pi R^3 n_c}{GI_t} F \left(\frac{\sin^2\beta I_t}{\hat{\mu} A R^2} + \frac{\cos\beta I_t}{AR^2} + \cos\beta + \frac{\sin^2\beta I_t}{\hat{\mu} I_\zeta} \right). \tag{9.31}$$

Neglecting the terms corresponding to $F_{ax,t}$, $F_{ax,b}$, and $M_{ax,b}$, Eq. (9.31) turns into the commonly used approximation formula for axial loaded coil springs (cp. [83,153]). It is tolerable to consider only terms associated with $M_{ax,t}$ under two assumptions:

- widely wound wire (a,$b << R$)
- small pitch ($\beta << 1$) .

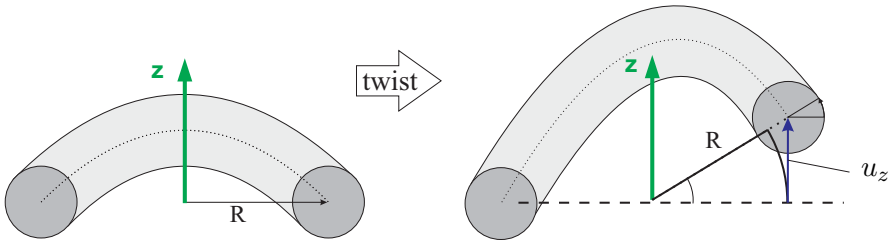


Figure 9.9: Geometric coupling between rotation and translation

Under these assumptions the elastic energy is almost completely stored in the torsional deformation of the cross section. According to Eq. (9.25) torsion is predominantly affected by ϑ_t. Under the assumptions

$$u_t = 0, \qquad u_n = 0, \qquad u_b = 0, \qquad \vartheta_n = 0, \qquad \text{and} \quad \vartheta_b = 0 \tag{9.32}$$

for an axial loaded spring, the system of PDEs in Eq. (9.27) can be reduced:

$$GI_t \frac{\partial^2 \vartheta_t}{\partial s^2} + d_t = \rho A R^2 \frac{\partial^2 \vartheta_t}{\partial t^2}\,. \tag{9.33}$$

As the spring is loaded axially, a twist of the individual cross sections will produce a displacement u_z of the spring wire in $\boldsymbol{z}$-direction. Figure 9.9 illustrates this relationship. Under the assumption of small rotations, $u_z \propto R\vartheta_t$ is a good approximation and Eq. (9.33) is transformed into Eq. (9.34).

$$\rho A \frac{\partial^2 u_z}{\partial t^2} - \frac{GJ}{R^2} \frac{\partial^2 u_z}{\partial s^2} = f_z \tag{9.34}$$

This is a hyperbolic PDE, also known as the one-dimensional *wave equation*, which can be solved analytically for the homogeneous case.

9.3.5 Validation of the Wave Equation

For an axial loaded spring, the general system of partial differential equations has been reduced to one degree of freedom using the assumptions of widely wound spring wire and small pitch.

To validate these, a detailed FEM model of a coil spring has been built for static and dynamic calculations. The results for differing geometries (curvature, pitch, and ellipticity) and boundary conditions are compared to the analytic solution of Eq. (9.34). In the static case deformations are compared whereas in the dynamic case natural frequencies are matched.

The FEM analysis has been done with ANSYS®Academic Research Vers. 11.0. The spring is discretised with approximately 3000 SOLID92-elements. The reference set

Table 9.1: Acceptable parameters for 1% and 2% error limit

variation	error $< 1\%$	error $< 2\%$
curvature R/a	< 2.6	< 2.3
pitch β	$< 17\,^\circ$	$< 21^\circ$
ellipticity a/b	< 1.3	< 1.5

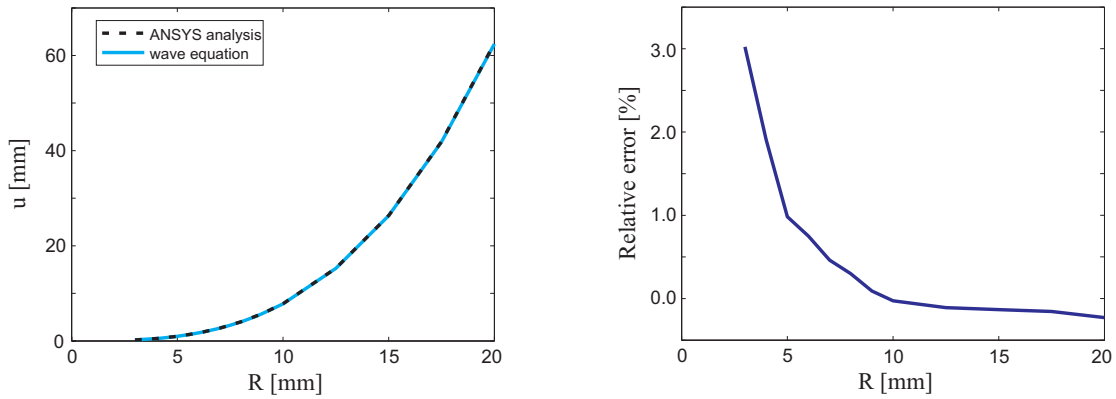
of parameters for a coil spring with $n_c = 8$ active coils is

$$R = 10\text{mm}, \qquad h = 5\text{mm}, \qquad a = b = 1.5\text{mm}$$
$$E = 210\,\text{GPa}, \qquad \mu = 0.28, \qquad \rho = 7800\,\frac{kg}{m^3}\,. \tag{9.35}$$

Before discussing the results in detail, an overview is given in Table 9.1. The relative error is expected to be smaller than 1% or 2% for the given variations. The assumptions seem to be valid for commonly used valve springs.

Static Analysis

Curvature The coil radius R is varied while the cross section and pitch are kept constant to investigate the influence of the curvature. Assumptions concerning the radius have been made at two passages: First, curvature is assumed to be small to derive the six coupled PDEs. Second, $\frac{J}{AR^2}$ is neglected in order to reduce the full PDEs to the one-dimensional wave equation. The validation of both assumptions in total is shown in Figure 9.10. The maximum deviation is approximately 3%.

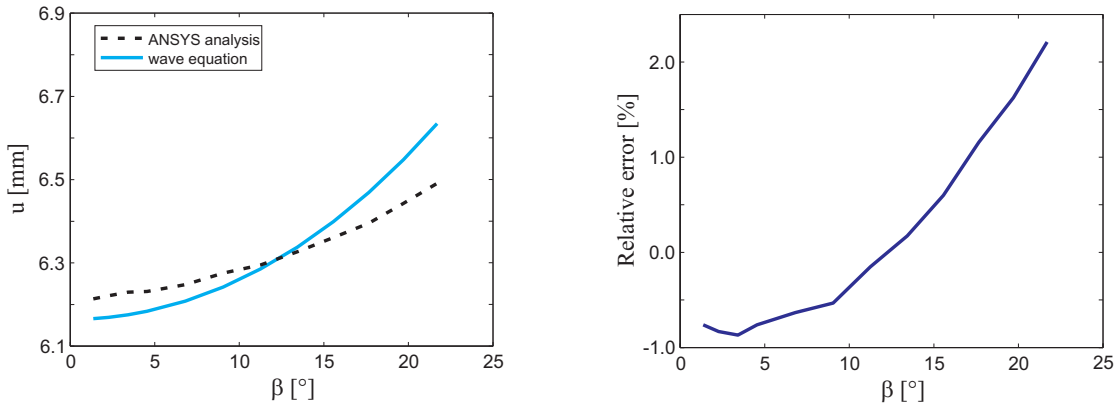


(a) Variation of radius R: Displacement

(b) Relative error.

Figure 9.10: Effect of curvature

Pitch Next, the pitch of the spring is varied while keeping the remaining parameters R, a, and b constant. The stiffness is overestimated for a small pitch and underestimated for a large pitch (see Figure 9.11). The maximum error occurs for high pitches and is approximately 2%. For high values of β, torsion is not the only

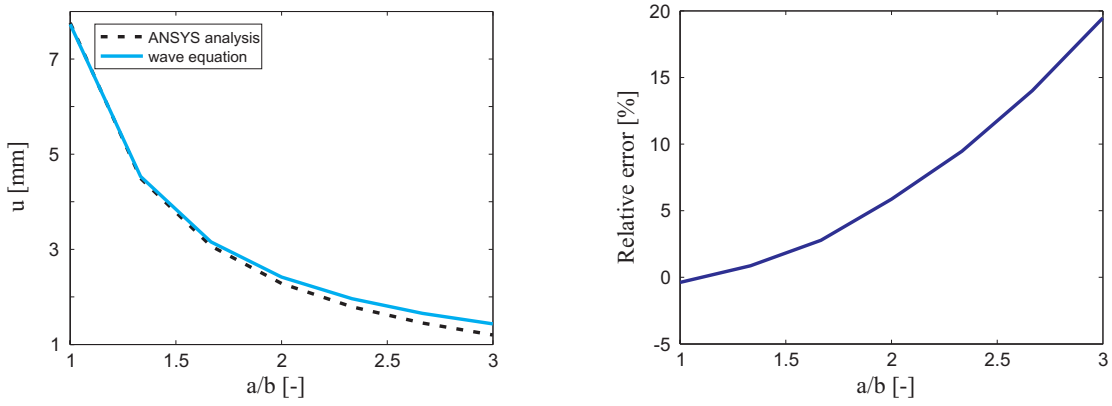


(a) Variation pitch β: Displacement (b) Relative error

Figure 9.11: Effect of pitch

resultant reaction due to axial load, as assumed in the derived model. Hence, it is too flexible.

Ellipticity To prove the assumptions regarding warping, the cross sectional shape is varied from circular to elliptical. Warping does not occur for a straight rod with a circular cross section. For explicit elliptical shapes the stiffness of the spring is underestimated (see Figure 9.12). Compared to the variation of R and β, the relative error is rather high. For a ratio $\frac{a}{b} = 3$, the deviation is about 20%.



(a) Variation a β: Displacement (b) Relative error

Figure 9.12: Effect of ellipticity

Dynamic analysis

For dynamic analysis the natural frequencies of the ANSYS FEM model are compared to those of the wave equation. As each node of the ANSYS model has three degrees of freedom, there are more than one dominant motion directions within the eigenmodes. Distinct dominant motions such as lateral, extensional, and torsional

are identifiable. According to Table 9.2, the first four extensional modes are relatively accurate. As the dynamics of the spring are described by the wave equation, the natural frequencies are increased by an integral multiple of the wave propagation speed divided by the wave length. As can be seen in Table 9.2, all frequencies are separated by approximately 341 Hz. The extensional modes of the ANSYS analysis are not separated equidistantly. Results from [130] indicate that by considering motion in all directions, higher frequencies are obtained more accurately.

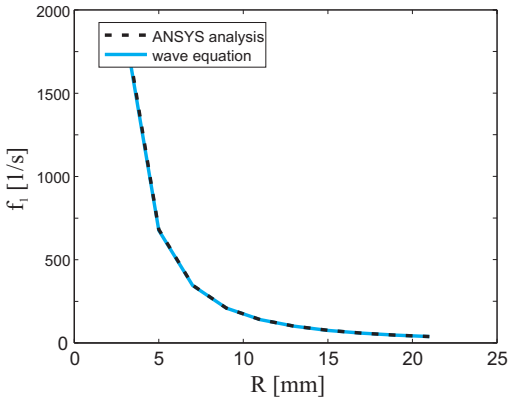
Table 9.2: First 16 natural frequencies for a spring with clamped-free boundary conditions

number	ANSYS [Hz]	wave equation [Hz]	primary motion
1	90.02		lateral
2	90.12		lateral
3	171.2	170.5	extensional
4	191.1		torsional
5	363.0		lateral
6	369.7		lateral
7	509.4	511.6	extensional
8	567.7		torsional
9	752.2		lateral
10	766.4		lateral
11	840.1	852.7	extensional
12	924.0		torsional
13	1024		lateral
14	1033		lateral
15	1116	1194	extensional
16	1226		torsional

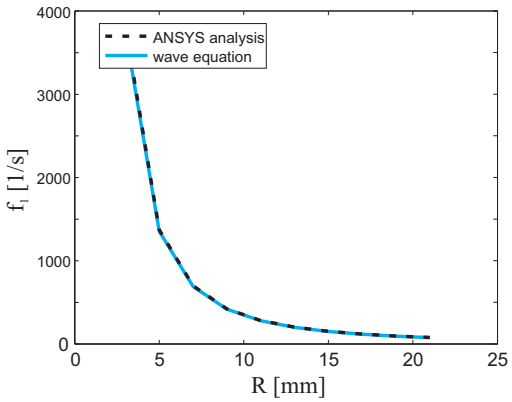
Next, the same parameter variations as for the static case are carried out calculating the natural frequencies. As a measure of the accuracy of the assumptions, only the first extensional natural frequency is compared. Two different cases are investigated. First, the spring is clamped at one end and free at the other end. Second, both ends of the spring are clamped. The results show the same trend as in the static case (see Figures 9.13, 9.14, and 9.15).

9.3.6 Coil Springs as Multi-Body Systems

The one dimensional partial differential equation derived in the previous sections may be solved in closed form for the homogeneous case. As for practical applications coil clashing, changing boundary conditions, and arbitrary excitation have to be considered, and a finite element formulation augmented with coil-to-coil contacts is necessary. Therefore, the wave equation is discretised using standard FE methods (e.g. [23] or [204]) with linear and quadratic Lagrange, as well as cubic Hermite polynomials as form functions to yield Eq. (9.36) [107].

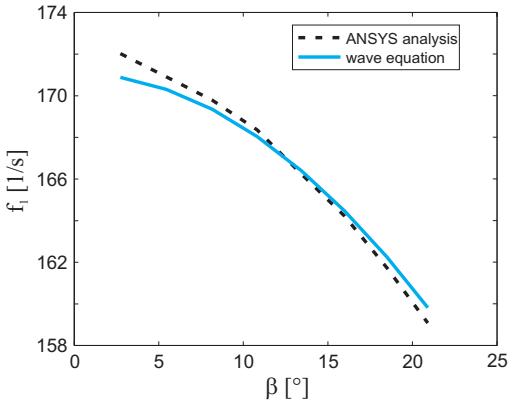


(a) Clamped-free boundary conditions

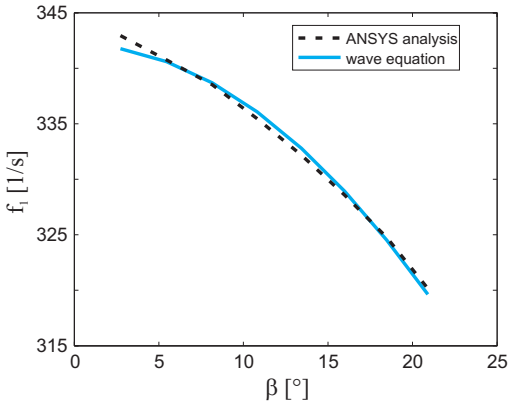


(b) Clamped-clamped boundary conditions

Figure 9.13: Effect of curvature

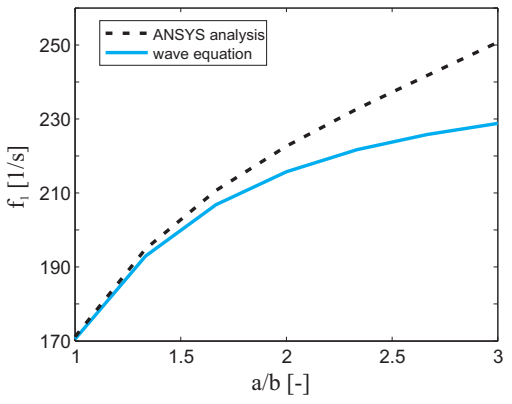


(a) Clamped-free boundary conditions

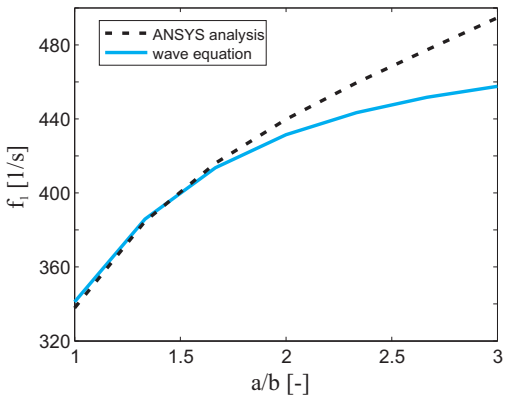


(b) Clamped-clamped boundary conditions

Figure 9.14: Effect of pitch



(a) Clamped-free boundary conditions



(b) Clamped-clamped boundary conditions

Figure 9.15: Effect of ellipticity

Damping

To take damping into account three different damping approaches with damping matrix $\boldsymbol{D}$ according to Eq. (9.37) are applied to the coil spring. The equations of motion are extended with a velocity proportional damping and result in Eq. (9.36).

$$\boldsymbol{M}\ddot{\boldsymbol{q}} + \boldsymbol{D}\dot{\boldsymbol{q}} + \boldsymbol{K}\boldsymbol{q} = \boldsymbol{F} \tag{9.36}$$

$$\boldsymbol{D} = \boldsymbol{D}_1 + \boldsymbol{D}_2 + \boldsymbol{D}_3 \tag{9.37}$$

The damping matrix $\boldsymbol{D}$ is composed of three effects: Rayleigh damping due to material and air damping ($\boldsymbol{D}_1$), discrete dampers between the degrees of freedom ($\boldsymbol{D}_2$), and damping applied to non-active coils ($\boldsymbol{D}_3$). The last type of damping is motivated by the energy dissipation due to friction between rubbing coils in contact. If coils clash or come into contact, the coils will rotate against each other or perform small movements in the contact plane as the pitch changes due to compression or expansion. This kind of motion has not been considered in the derived coil spring model. Additional dampers are added to the degrees of freedom if the corresponding coils are in contact.

9.4 Test Rig

A test rig was designed and built to obtain experimental data for validating of the spring model. An overview of the test rig with its main functionality is given in Figure 9.16. More details can be found in [34].
The basic principle of the test rig is that a roller is pressed against a cam with a stiff spring so that the roller follows the cam contour. With this construction the rotational motion of the cam is transformed in a translational motion of the cam follower. The cam follower is mounted with a hydrostatic bearing to minimise friction and wear and to make high revolution speeds possible (Figure 9.17).

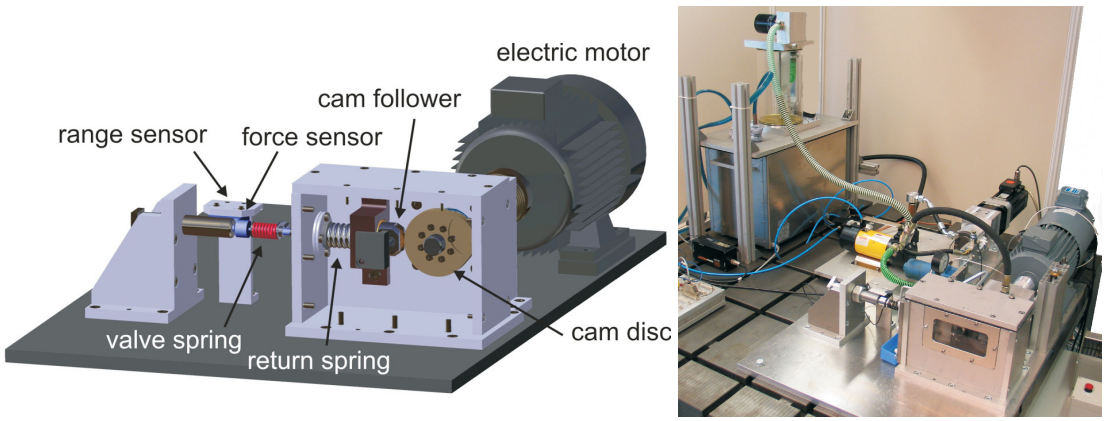


Figure 9.16: Test rig for valve springs

The test rig is designed to stimulate springs with different types of excitation. Mainly two different excitations are used to validate the spring model: cam discs with a harmonic excitation and cam discs with an excitation similar to the way a valve travels in a combustion engine. At the clamped side a force sensor based on the strain gauge technique is installed. At the free side a laser sensor for large displacements or an eddy current sensor for small displacements is used to measure the position of the rod (Figure 9.18).

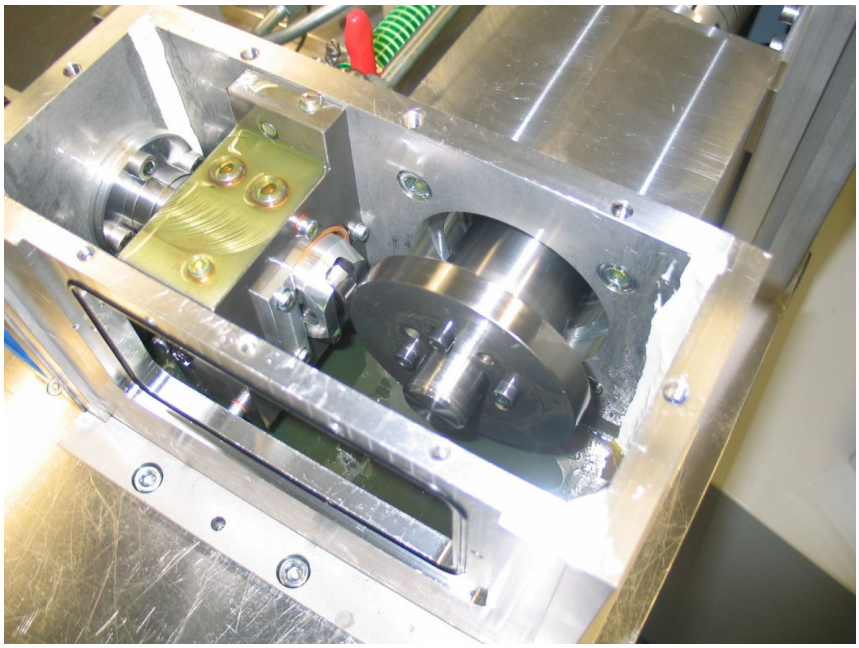


Figure 9.17: Details of the test rig: cam disc, cam follower and hydrostatic bearing

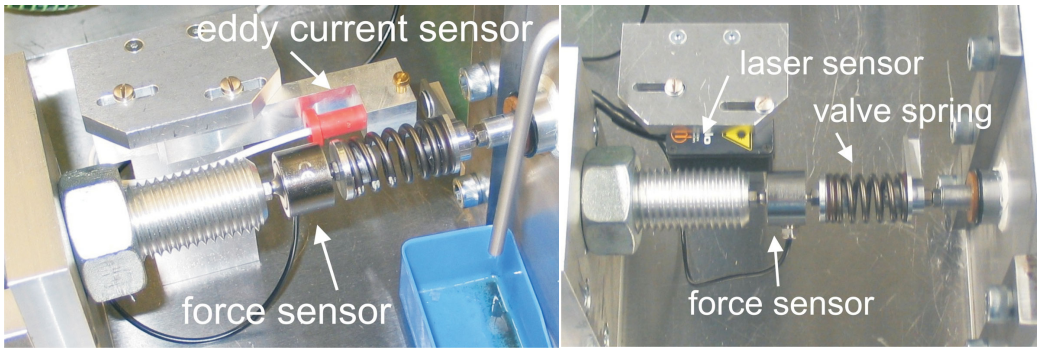


Figure 9.18: Details of the test rig: laser sensor, eddy current sensor and force sensor

9.5 Results

In Section 9.3.5 it was shown that the wave equation is an appropriate approximation for coil springs under axial load.

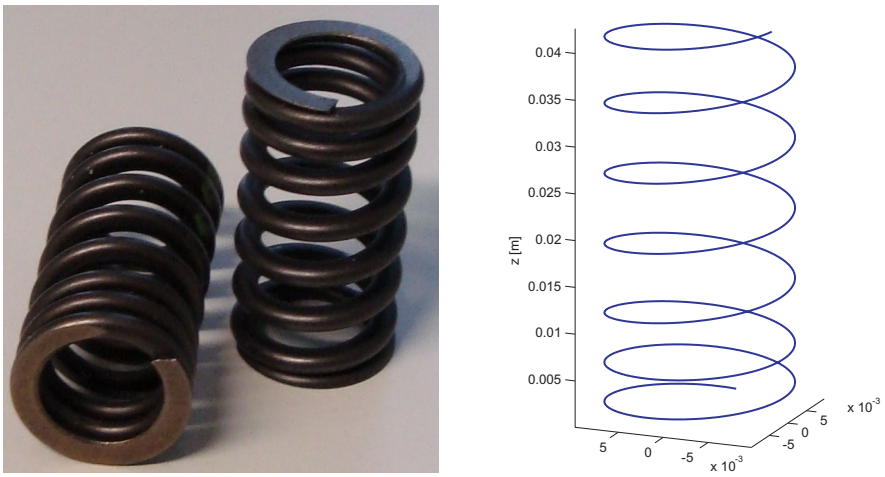


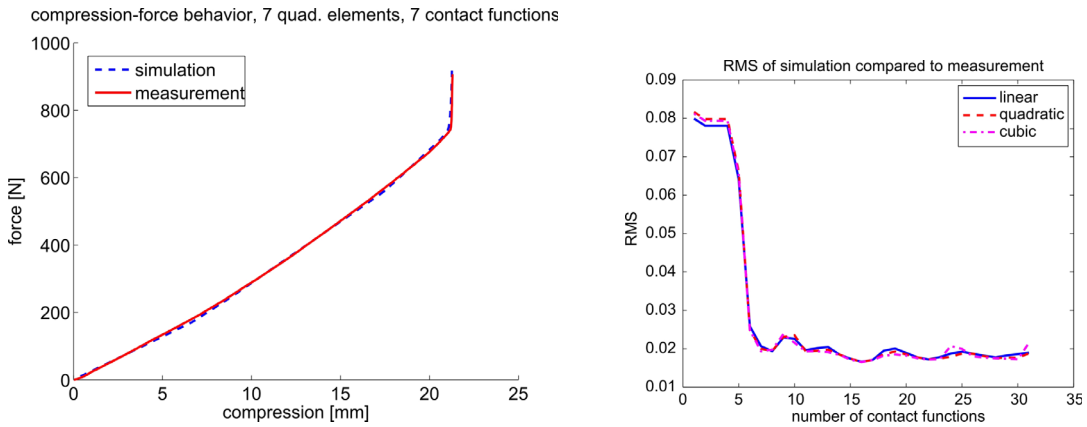
Figure 9.19: Valve spring from combustion engine

Next, the results of the multi-body model are compared to experimental data. This is done by validating the static characteristics as well as validating the dynamics in the frequency and time domain.
The subject of investigation is a valve spring of a current combustion engine having an elliptical cross section (2a= 3.8mm, 2b= 3.1mm), progressive behaviour, and seven active coils (Figure 9.19).

9.5.1 Static Characteristics

For static validation the model is verified using load displacement characteristics. Figure 9.20(a) shows the load displacement of the spring with about one quadratic element per coil and seven coil to coil contacts. A very good correspondence between simulation and measurement can be observed, even for large compressions.

In Figure 9.20(b) the error in terms of the RMS value is plotted against the number of contacts for different elements with comparable numbers of degrees of freedom: 22 linear elements (23 dofs), 11 quadratic elements (23 dofs), and 10 cubic elements (22 dofs) were used.
Because the number of contacts has a serious influence on the numerical effort, only as many contacts as necessary should be used. According to Figure 9.20(b), about one contact per coil element seems to be sufficient for this type of valve spring in the static case.



(a) Static behaviour: progressive characteristic (b) RMS simulation vs. measurement

Figure 9.20: Static validation of the progressive characteristic and error as a function of the number of contacts

9.5.2 Validation of the Dynamics

For validating the dynamics, measurements from the cam disc test rig are used. In the frequency domain, the spring is excited with a harmonic excitation of 0.05mm amplitude and the eigenfrequencies are compared. In the time domain, the excitation is comparable to the valve lift of the corresponding combustion engine. For this purpose, an appropriate cam disc has been designed. The measured kinematic quantities are used as kinematic excitations for the simulation model. The measured force is compared to the force obtained by the simulation model. A spring model with seven contacts and seven quadratic elements is used.

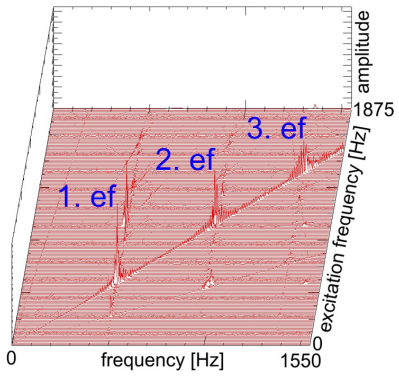
Frequency Domain

For different pre-stresses, the spring is harmonically excited in a quasi-static run-up. Analysing the measured data with Fourier transformation, a waterfall diagram as in Figure 9.21(a) is obtained for every discrete pre-stress. As the amplitudes are small, the obtained frequencies can be interpreted as pseudo-eigenfrequencies for the particular configuration.

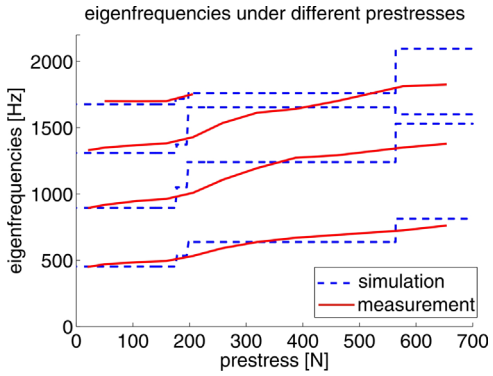
The comparison between the simulated and measured pseudo-eigenfrequencies of the valve spring is shown in Figure 9.21(b). It must be considered that the simulation model can only produce discrete eigenfrequencies because the contacts close at discrete points. Nevertheless, the deviation between the measured and simulated eigenfrequencies is small.

Time Domain

The measured time responses are used to find appropriate values for the damping parameters. In this work three different approaches for energy dissipation are used:

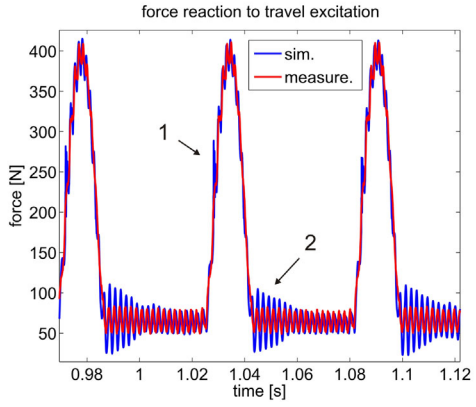


(a) Waterfall diagram of a run-up

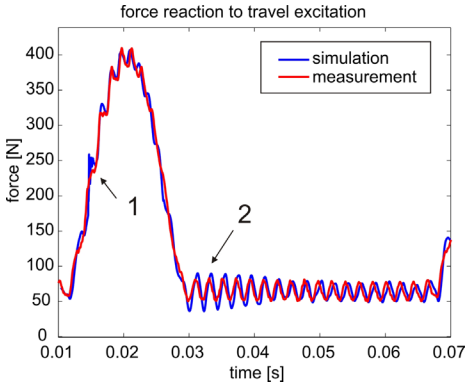


(b) Pseudo eigenfrequencies under different load

Figure 9.21: Frequency domain



(a) Material and discrete damping



(b) Final results

Figure 9.22: Force at the clamped end of an excited spring: different damping approaches

material damping, energy dissipation due to rubbing coils, and damping in the contacts between coils.
Dissipation at contact points depends on the contact model in use. For rigid contact formulations, impacts are considered to be fully plastic and therefore no damping parameters are needed. For flexible contacts, viscose damping is applied in the contacts. To take into account energy dissipation due to rubbing coils, discrete dampers are used that are only active for coils being in contact.

The effects of different damping approaches can be seen in Figure 9.22. In Figure 9.22(a) material damping according to Rayleigh and discrete dampers are considered. The simulated force diverges from the measured force at two points: At point 1 a contact closes and therefore creates an overshooting of the force. At point 2 contacts are reopened, which also lead to an overshooting of the force. Apart from this, the simulation and measurement are consistent. Especially the frequencies are in good correlation. This result suggests that too much energy is stored in the contacts and energy dissipation is too small.

Based on this consideration, discrete dampers are used for degrees of freedoms of coils being in contact to model the dissipation due to rubbing coils. The corresponding damping matrix has to be recalculated at every time step. The results are shown in Figure 9.22(b).
The effect of overshooting forces can also be decreased slightly by using rigid coil to coil contacts with set-valued force laws instead of flexible contacts. Within rigid contact formulations impacts are considered to be fully plastic and therefore maximum energy is dissipated in the contact zone.

10 Hydraulic Lash Adjuster

10.1 General Introduction to Hydraulic Lash Adjusters

The hydraulic lash adjuster (HLA), also known as hydraulic valve lifter, is a standard construction element of modern valve trains. Valve trains without self adjusting valve lifters require a small clearance between rocker arm and valve to allow thermal expansion, valve lengthening, and abrasive wear. However, if the clearance is too large, this will result in excessive noise and component wear. While if the clearance is too small or negative, the valve will not be able to fully close causing pressure losses in the cylinder and burnt valves.
It is possible to operate the valve train with zero clearance using a hydraulic lash adjuster. Thus, the cam surface and entire mechanism is less subject to mechanical stress and the operating noise is reduced. Maintenance and assembly costs are reduced as there is no longer the need to manually adjust the valve clearance. Furthermore, more precise valve timing for improved thermodynamic processes is possible.
Besides hydraulic lash adjusters, there are also mechanical lash adjusters e.g. with buttress threads and return springs [136]. Lifters with expanded functionality are used for cylinder deactivation [103]. Deactivation mechanisms within the lash adjuster are actuated by solenoids to release lost motion components and allow the lash adjuster to telescope. The pivot function is deactivated, the lifters are unable to support the rocker arm, and the valves remain closed [7].
Nevertheless, there are also some potential problems with hydraulic lash adjusters. Oil draining from the lifters, air inside the high pressure chamber, or blocked oil feed are possible sources for tappet noise and rattling noise at engine start-up. Under certain circumstances, a HLA can pump up creating a negative valve clearance so that its valve cannot close. Lifter pump-up is serious, as there may be the risk of a collision between valve and piston or burnt valves.

10.1.1 Design and Function

As shown in Figure 10.1 the hydraulic lash adjuster consists of a hollow plunger (1), a check valve (2) with ball and spring, and a plunger spring (3) between plunger and housing (4) that lifts the plunger even without oil pressure. The volume between the plunger and housing is called the high pressure chamber. It is connected to the low pressure or supply chamber via an orifice with check valve. The lash adjuster

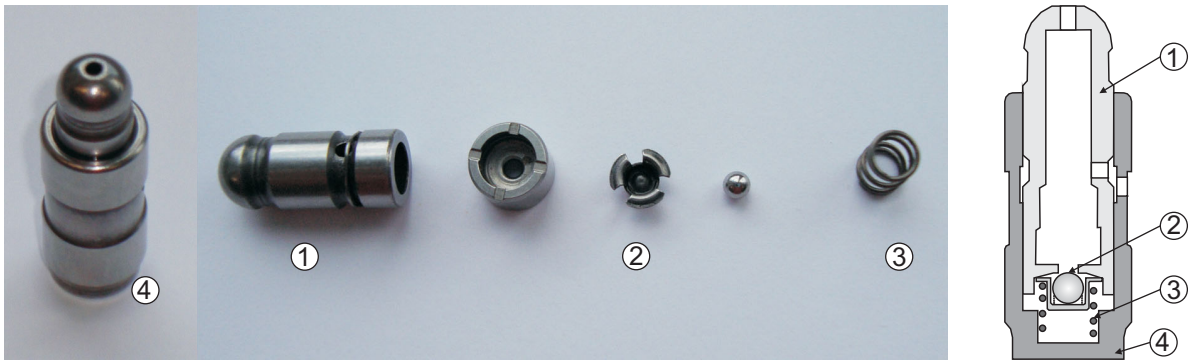


Figure 10.1: Components of the hydraulic lash adjuster

is supplied with engine oil from an oil gallery through a small drilling. The lifter is located between valve and camshaft at the pivot point of the roller cam follower. During the base-circle phase, the valve is closed and the lash adjuster is free to fill with oil. The return spring and the hydraulic pressure from the supply line forces the plunger against the finger follower removing any existing gaps between the valve train components. When the valve is opening, compressive forces are applied to the lifter by the camshaft. The plunger is forced into the hosing, the check valve closes and the pressure in the high pressure chamber increases. Due to the pressure difference between the high pressure chamber and supply line, oil is squeezed through the leakage gap. This is the only form of damping inside the HLA. At the end of the valve lift, the compressive forces on the plunger become smaller than the hydraulic forces and return spring forces on the plunger. Thus, the high pressure chamber expands to occupy any clearance. The check valve opens and the leaked oil is replenished.

10.1.2 Literature Survey

Compared to a solid supporting point for the rocker arm, a hydraulic lash adjuster induces additional flexibility in the valve train and has a non-negligible influence on the overall dynamics of the valve train. A deep insight into HLA dynamics is required, making detailed simulation models and experimental analyses necessary.
In the context of the simulation of timing assemblies many publications exist regarding lash adjusters. Kreuter and Maas [121, 135] discussed the influence of the lifter on valve bouncing utilising measurements of the plunger motion.
Bakaj [12] was among the first to use simulation techniques to analyse the dynamics of valve trains including hydraulic lash adjusters. Philips et al. [163] included a reduced model of the HLA in their valve train model. This model was based on a spring damper model characterised by experimental data including leakage and the influence of air on oil stiffness.
In the earlier works of Pisano, a simplified hydraulic model neglecting oil stiffness is presented [32]. Later, this model was amended with the influence of the air content on oil stiffness [93].

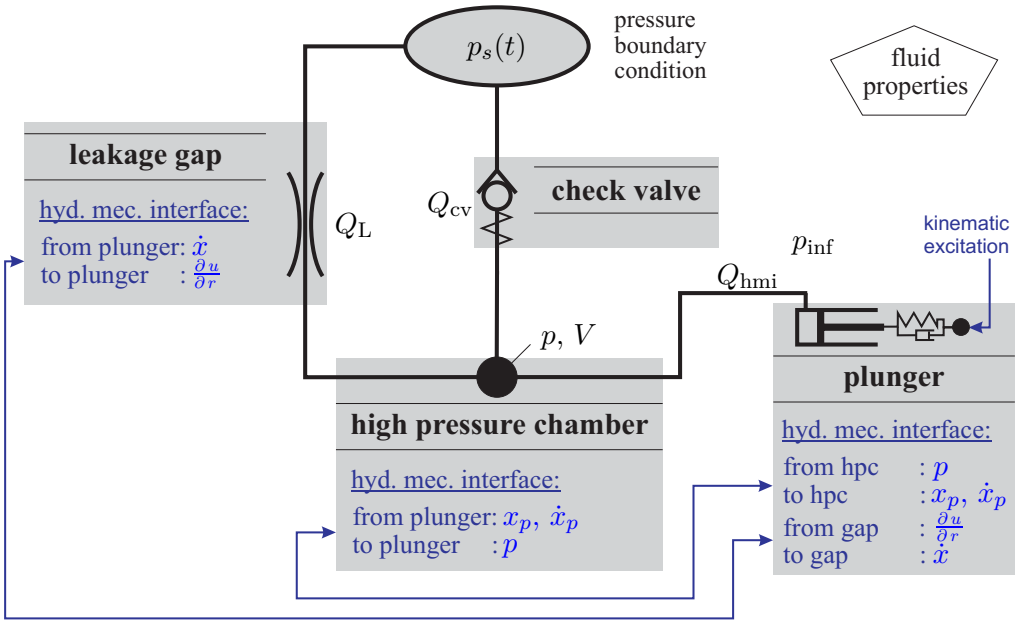


Figure 10.2: Simulation model of the hydraulic circuit and the coupling between hydraulics and mechanics

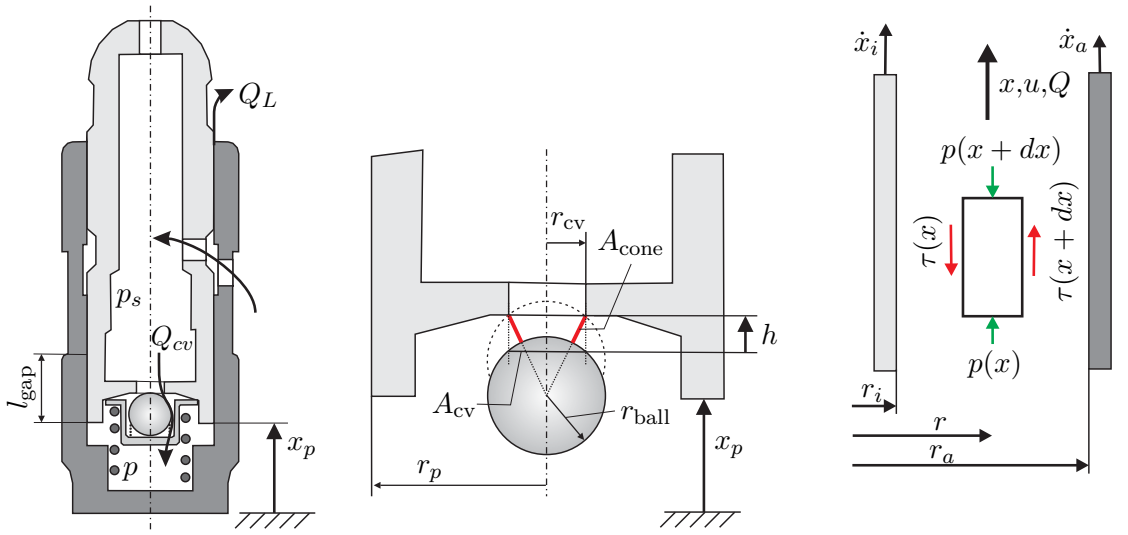
Zou and McCormick [207] developed a hydraulic model of the lash adjuster including the check valve. Porot and Trapy [167] applied a similar model to a direct acting valve train doing numerical and experimental studies. Further investigations on the dynamics of timing assemblies were done by Heiduk [95].

The design of basic hydraulic chain tensioners is similar to hydraulic lash adjusters. Both systems comprise standard hydraulic elements such as pressure chamber, check valve, and leakage gap. Therefore, reference is made to the comprehensive literature regarding chain tensioner at this point. Post [168] et al. investigated the influence of chain tensioners on the dynamics of timing chains. This work has been continued by Hösl [106] and Engelhardt [50] who derived simulation models based on characteristic diagrams obtained from CFD simulations. CFD-based modelling of chain tensioners has been refined by Krüger [123], showing an excellent correlation between experiment and simulation [124, 125]. Nicola [152] discussed the dynamic behaviour of chain drives and chain tensioners under rotational irregularity excitation with the help of simulation and experiment.

The lash adjuster model presented in this work extends the models known from literature by two aspects: First, the calculation of the drag coefficient is refined and second, the influence of the plunger elasticity on the leakage flow is considered.

10.2 Simulation Model

In the simulation framework, hydraulic elements are classified in nodes and links. Links serve as connectors between hydraulic volumes (nodes) or as an interface be-



(a) Geometrical data (b) Leakage gap

Figure 10.3: Model of the hydraulic lash adjuster

tween hydraulic and mechanical components. Modelling hydraulic systems means to apply this logical structure and to describe the real system in terms of links and nodes.
The hydraulic circuit of the HLA is depicted in Figure 10.2. The model comprises the high pressure chamber as hydraulic node with variable volume, as well as a check valve and a leakage gap as hydraulic link.
The interaction between both physical domains—hydraulics and mechanics—is attained by the plunger. The motion of the plunger changes the volume of the high pressure chamber and the flow condition inside the leakage gap. Vice versa, the pressure p inside the hydraulic node and the velocity of flow u inside the gap result in forces on the plunger. The coupling between the mechanical and hydraulic system equations is illustrated by the interface quantities and arrows in blue. As the dynamics of the check valve ball is only influenced by its own hydraulic quantities, the hydro-mechanical interface (hmi) for the ball is omitted in this drawing.
Furthermore, the fluid properties and the boundary conditions in terms of plunger excitation and hydraulic pressure p_s in the supply line have to be taken into account. In the next section, the governing equations describing the hydraulic behaviour of the hydraulic components are derived. Mechanical modelling, e.g. spring forces or contact forces between ball and check valve, is done as described in Chapter 2.

10.2.1 High Pressure Chamber

To a certain degree oil circulating in a combustion engine is always mixed with air. Hence, the compressibility of hydraulic fluids is not a priori negligible. Considering compressibility, the high pressure chamber is described as an elastic hydraulic node with pressure p and volume V. The rate of pressure change $\dot{p}$ is proportional to

the hydraulic (Q_{cv}, Q_L) and mechanical flow rate (Q_{hmi}) with bulk modulus K_{fl} as proportionality factor. K_{fl} is evaluated with respect to the air ratio of the oil.

$$\dot{p} = \frac{K_{\mathrm{fl}}}{V} \left(Q_{\mathrm{cv}} - Q_L + Q_{\mathrm{hmi}} \right) \tag{10.1}$$

$$V = V_0 + V_{\mathrm{hmi}} \tag{10.2}$$

The mechanical flow rate Q_{hmi} is caused by the motion of the plunger. The current volume V is composed of the initial volume V_0 and the volume V_{hmi} due to the plunger displacement and is obtained by integrating Q_{hmi}.

$$Q_{\mathrm{hmi}} = -\dot{x}_p r_p^2 \pi \tag{10.3}$$

$$V_{\mathrm{hmi}} = r_p^2 \pi \left(x_p - x_{p,0} \right) \tag{10.4}$$

Expressions for the flow rate Q_{cv} through the check valve and through the leakage gap (Q_L) are derived in the next sections.

During valve lift operation, vertical forces are supported by the oil volume of the high pressure chamber and transferred by the plunger. The equations of motion for the plunger result from mechanical forces due to limit stops, the plunger spring, and the connection to the roller cam follower, as well as from hydraulic forces $F_{\mathrm{hmi,plunger}}$ of the hydraulic mechanical interconnections.

$$F_{\mathrm{hmi,plunger}} = \left(p - p_{\mathrm{inf}} \right) r_p^2 \pi + r_{\mathrm{cv}}^2 \pi (p_s - p) + \eta \left. \frac{\partial u}{\partial r} \right|_{r=r_i} \tag{10.5}$$

The last summand in Eq. (10.5) arises from friction due to leakage flow through the gap between piston and housing. It represents the main contribution to system damping of the lash adjuster.

10.2.2 Check Valve

The check valve is used to prevent the oil from flowing back into the supply chamber and the plunger from sinking down during valve operation. Fig. 10.3(a) shows the check valve with orifice and spring-loaded ball. In many cases it is not sufficient to model only the unilateral characteristic of the check valve. It is also necessary to consider the dynamic behaviour of the ball and hydraulic losses. In general, the pressure loss at given flow velocity v is characterised by the resistance coefficient ζ according to Eq. (10.6) [200]:

$$\Delta p = \frac{1}{2} \rho v^2 \zeta \, . \tag{10.6}$$

The main loss in pressure occurs while opening the check valve at the orifice and at the annular section by the released ball. The calculation of the resistance coefficient becomes complicated as the pressure loss is a function of the Reynolds number and ball position. The modelling of the pressure losses is one of the main parameters

that determine the sophistication of the lash adjuster model.
A very detailed model of a check valve within a chain tensioner is presented by Engelhardt [50]. It is based on resistance coefficients as a function of the Reynolds number Re and the ball position obtained by CFD simulation. For the modelling of hydraulic lash adjusters within valve train simulation, such a model depth is normally not required. Therefore, the relation between the volumetric flow rate Q and the pressure loss Δp is derived starting with the orifice formula [55]:

$$Q_{cv} = \operatorname{sgn}(\Delta p)\, \alpha_{\text{eff}}\, A_{cv} \sqrt{\frac{2\,|\Delta p|}{\rho}}\,, \tag{10.7a}$$

$$\Delta p = p_s - p\,, \tag{10.7b}$$

$$A_{\text{cv}} = r_{cv}^2\, \pi\,. \tag{10.7c}$$

The orifice or flow factor α_{eff} summarises losses due to all mechanisms including the influence of the cross section ratio, the contraction coefficient, and friction. In general, α_{eff} is a function of the ball position and Re.
Comparing Eq. (10.7a) to Eq. (10.6), a relation can be derived between the flow factor α and the resistance coefficient ζ:

$$\alpha_{\text{eff}}^2 = \frac{1}{\zeta} = \frac{\frac{1}{2}\rho v^2}{\Delta p}\,. \tag{10.8}$$

Determining the losses is one of the main challenges of modelling a check valve. If the influence of the Reynolds number is neglected, one possibility is to use the resistance coefficient given by Idelchik [111]:

$$\zeta \;=\; 2.7 - \beta_1 + \beta_2 \tag{10.9}$$

with:

$$\beta_1 \;=\; 1.6\,\frac{r_{cv}}{h}\,,$$

$$\beta_2 \;=\; 0.56\left(\frac{r_{cv}}{h}\right)^2\,.$$

As Eq. (10.9) is only valid for $0.2 < \frac{h}{r_{cv}} < 0.5$, another approach is given by Eq. (10.10). To model the ball position dependency, the area A_{cone} of a truncated cone is used as flow cross section area (cp. Figure 10.3(a)) in combination with a constant flow coefficient α. This approach is implemented for example in the ADAMS/Engine software package [148].

$$Q_{\text{cv}} = \operatorname{sgn}(\Delta p)\, \alpha\, A_{\text{cone}} \sqrt{\frac{2\,|\Delta p|}{\rho}} \tag{10.10a}$$

$$A_{\text{cone}} = \pi\,(\ell^2 - r_{\text{ball}}^2)\frac{r_{\text{cv}}}{\ell} \tag{10.10b}$$

$$\ell = \sqrt{(h + \sqrt{r_{\text{ball}}^2 - r_{\text{cv}}^2}\,)^2 + r_{\text{cv}}^2} \tag{10.10c}$$

The flow rate Q_{cv} is equal to zero if the ball is in contact with the orifice ($h = 0$). The boundary value of the flow rate for large values of h goes to infinity according to Eq. (10.10). But from a physical point of view, one would expect the flow rate to be independent from the ball position for high values of h and having the limit

$$\lim_{h\to\infty} Q_{\mathrm{cv}} = \mathrm{sgn}(\Delta p)\,\alpha\,A_{\mathrm{cv}}\sqrt{\frac{2\,|\Delta p|}{\rho}}\,. \tag{10.11}$$

To describe the flow resistance in a physically correct way even at the limits, a new approach is used within this thesis. The reformulated flow equation Eq. (10.12) converges for small values to Eq. (10.10) and for large values to Eq. (10.11):

$$Q_{\mathrm{cv}} = \mathrm{sgn}(\Delta p)\,\alpha\sqrt{\frac{2\,|\Delta p|}{\rho}}\left(A_{\mathrm{cone}}\left(\frac{r_{\mathrm{ball}}}{\ell}\right)^2 + A_{\mathrm{cv}}\frac{\ell - r_{\mathrm{ball}}}{\ell}\right) \tag{10.12}$$

The effective flow coefficient for all three approaches is illustrated in Figure 10.4. Furthermore, the allowed positions of the ball between inner and outer stop are drawn. For small values of h, all three flow coefficients are in good agreement. For larger h values, both the approach according to Idelchik and the derived new approach tend to a constant value, whereas the approach used within ADAMS/Engine yields unbounded flow rates.

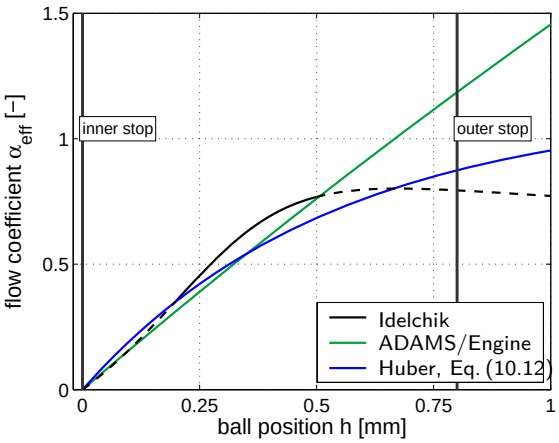


Figure 10.4: Effective flow coefficient check valve for different ball positions

The formulas derived in the previous section describe the flow rate as a function of the pressure difference neglecting inertia forces of the fluid. To consider fluid inertia Eq. (10.13) has to be added.

$$\dot{Q}_{\mathrm{cv}} = \frac{A_{\mathrm{cv}}}{\rho\ell_{\mathrm{cv}}}\left(p_s - p - \Delta p_{\mathrm{loss}}\right) \tag{10.13}$$

The pressure loss Δp_{loss} is equivalent to Δp of the previous orifice formulas and can be expressed as a function of Q_{cv} yielding a differential equation for the flow rate.

For a simplified check valve model, the position of the check valve ball can be calculated from the static equilibrium of pressure and spring forces neglecting inertia and damping forces. As these forces are essential for the check valve dynamics, the ball position is considered as a mechanical degree of freedom within this thesis. The equations of motion contain mechanical forces due to the spring and due to the

contact between ball and plunger, as well as hydraulic forces $F_{\text{ball,hmi}}$:

$$F_{\text{ball,hmi}} = -A_{\text{cv}}\,\Delta p - F_\eta\,. \tag{10.14}$$

The hydraulic force F_η caused by the oil viscosity can be interpreted as a damping force and is given by:

$$F_\eta = \frac{1}{2}\rho v_{\text{rel}}^2 c_w r_{\text{ball}}^2 \pi \qquad \text{with} \tag{10.15a}$$

$$v_{\text{rel}} = \frac{Q_{\text{cv}}}{A_{\text{cv}}} k_1 + \dot{h}\,. \tag{10.15b}$$

The parameter k_1 is first and foremost determined by the ratio $\frac{A_{\text{cv}}}{A_{p,i}}$ of cross sectional areas ($A_{p,i}$ is the inner cross section area of the plunger), but is also used to adjust the simplified flow conditions underlying Eq. (10.15). The drag coefficient is given by:

$$c_w = \begin{cases} k\frac{24}{\text{Re}} & \text{Re} \le \text{Re}_{\text{lim}} \\ 0.4 + k\frac{26}{\text{Re}^{0.8}} & \text{Re} > \text{Re}_{\text{lim}} \end{cases} \tag{10.16}$$

$$\text{with Reynolds number } \text{Re} = \frac{2\,v_{\text{rel}}\,r_{\text{ball}}\,\rho}{\eta}\,.$$

For small Re numbers the resistance force is calculated according to Stokes' law [190], whereas for large Re numbers Prandtl's approximation formula [46] supplemented by the parameter k is used. The change over Reynolds number Re_{lim} has to be calculated so that the drag coefficient is continuous.
The parameter k considers wall effects and other uncertainties. For a sphere with radius r in a cylindrical vessel with radius R and height H, the parameter k has been determined by Ladenburg to $k = (1 + 2.1\frac{r}{R})(1 + 3.3\frac{r}{H})$ [190].

10.2.3 Leakage Gap

The volume of oil passing through the annular gap between the plunger and housing has two main effects on the lash adjuster dynamics. First, the plunger is retracting during the cam lift event resulting in a reduced cam lift. Second, the leakage flow generates viscous friction forces and is the main source for damping within the hydraulic lash adjuster. Modelling approaches for the leakage flow can be found in Hösl [106] and Heiduk [95].
The flow Q_L through the leakage gap is derived using a cylindrical geometry with moving boundaries and constant oil density. Neglecting fluid velocities in circumferential direction, a two dimensional flow is assumed. With cylindrical coordinates the continuity equation is written as:

$$\frac{1}{r}\frac{\partial\,(r u_r)}{\partial\,r} + \frac{\partial\,u_x}{\partial\,x} = 0\,. \tag{10.17}$$

In fluid dynamics, the equations of momentum describing the motion of fluid substances are called Navier-Stokes equations. Without gravity forces the two-dimensional Navier-Stokes equations for Newtonian fluids are given by Eq. (10.18) [193]:

$$\rho\frac{\partial u_r}{\partial t} = -\frac{\partial p}{\partial r} + \eta\left[\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial u_r}{\partial r}\right) + \frac{\partial^2 u_r}{\partial x^2} - \frac{u_r}{r^2}\right] \tag{10.18a}$$

$$\rho\frac{\partial u_x}{\partial t} = -\frac{\partial p}{\partial x} + \eta\left[\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial u_x}{\partial r}\right) + \frac{\partial^2 u_x}{\partial x^2}\right]. \tag{10.18b}$$

Due to the small oil volume within the gap, inertia forces are neglected while steady state and laminar flow conditions are assumed. In this case with $u_x = u_x(x,r)$ and $u_r = 0$, Eq. (10.17) and Eq. (10.18) simplify to

$$\frac{\partial u_x}{\partial x} = 0\,, \quad \frac{\partial p}{\partial r} = 0\,, \quad \text{and} \quad -\frac{\partial p}{\partial x} + \eta\left[\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial u_x}{\partial r}\right) + \frac{\partial^2 u_x}{\partial x^2}\right] = 0\,. \tag{10.19}$$

Hence, the velocity profile is constant in x-direction ($u = u(r) = u_x(r)$) and the pressure $p = p(x)$ is constant in the radial direction. The leakage flow is subject to Eq. (10.20):

$$\frac{\mathrm{d}\,p}{\mathrm{d}\,x} - \eta\,\frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}\,r}\left(r\frac{\mathrm{d}\,u}{\mathrm{d}\,r}\right) = 0\,. \tag{10.20}$$

Eq. (10.20) can be interpreted as the equilibrium of pressure and viscosity forces on an infinitesimally small fluid volume as illustrated in Figure 10.3(b). Applying the equilibrium of forces and expanding $p(x + dx)$ and $\tau(x + dx)$ in Taylor series yields:

$$\frac{\mathrm{d}\,p}{\mathrm{d}\,x} + \frac{1}{r}\frac{\mathrm{d}\,(\tau\,r)}{\mathrm{d}\,r} = 0\,. \tag{10.21}$$

For a Newtonian fluid [193] the shear stress τ is given by the velocity gradient or strain rate multiplied with the dynamic fluid viscosity η:

$$\tau = -\eta\frac{\mathrm{d}\,u}{\mathrm{d}\,r}\,. \tag{10.22}$$

Eq. (10.21) in combination with Eq. (10.22) forms Eq. (10.20) which has been derived from Navier-Stokes equations. With the boundary conditions $u(r = r_i) = \dot{x}_i$ and $u(r = r_a) = \dot{x}_a$ at the inner and outer surface of the gap, the boundary value problem is solved by integration:

$$u = \dot{x}_i + (\dot{x}_a - \dot{x}_i)\,\frac{\ln\frac{r}{r_i}}{\ln\frac{r_a}{r_i}} + \frac{1}{4\eta}\frac{\mathrm{d}\,p}{\mathrm{d}\,x}\left(r^2 - r_i^2 + (r_i^2 - r_a^2)\,\frac{\ln\frac{r}{r_i}}{\ln\frac{r_a}{r_i}}\right). \tag{10.23}$$

The flow rate is calculated by the integral of the velocity on the cross section area of the gap yielding:

$$Q_L = \pi \left(\dot{x}_a r_a^2 - \dot{x}_i r_i^2 - (\dot{x}_a - \dot{x}_i) \frac{r_a^2 - r_i^2}{2 \ln \frac{r_a}{r_i}} \right) - \frac{\mathrm{d}\, p}{\mathrm{d}\, x} \frac{\pi}{8\, \eta} \left(r_a^4 - r_i^4 - \frac{(r_a^2 - r_i^2)^2}{\ln \frac{r_a}{r_i}} \right) . \tag{10.24}$$

The first term in Eq. (10.24) is the flow rate due to motion of the plunger $\dot{x}_i = \dot{x}_p$ and housing $\dot{x}_a$ and the second term is caused by the pressure gradient.
The flow rate is proportional to the third power of the gap distance (cp. Eq. (10.24) or [138]) and it is therefore very sensitive to variations of the gap distance. Compared to chain tensioners the gap distance in hydraulic lash adjusters is very small and in the range of a few micrometres. Hence, it is possible that a radial deformation of the plunger due to oil pressure in the high pressure chamber has an influence on the leakage flow rate. With Young's modulus E_p and wall thickness t_p of the plunger, the change of the plunger radius is given by:

$$\Delta r_i = k_3 \frac{-\Delta p\, r_i^2}{t_p\, E_p} . \tag{10.25}$$

As Eq. (10.25) is exactly valid for infinite pipes, the real plunger geometry has to be adjusted by the parameter k_3. The deformation in radial direction of a pipe is larger than the deformation of the vessel-like plunger. Hence, k_3 is less than one and has to be identified by comparing simulation and measurement.

The influence of plunger elasticity in radial direction is shown in Figure 10.5. The force acting on the plunger—obtained by simulation and measurement—is drawn in the upper part of the figure. The lower part shows the absolute value of the difference between measurement and simulation. Considering radial elasticity results in a higher force level during pressure peaks, as well as in better agreement between simulation and experiment.

10.2.4 Fluid

The oil viscosity has a significant influence on the leakage flow rate and the pressure level in the high pressure chamber. As illustrated in Figure 10.6(a), the oil viscosity is strongly temperature dependent. To calculate the viscosity as a function of temperature, a double logarithmic linearisation according to Ubbelohde-Walther [195,201] is used. For lubricating oils the viscosity-temperature behaviour is specified by DIN 51563 [41]. With given kinematic viscosities ν_1 and ν_2 at two test temperatures T_1 and T_2 (in degree Kelvin), the slope m of the viscosity-temperature function is:

$$m = \frac{W_1 - W_2}{\log T_2 - \log T_1} . \tag{10.26}$$

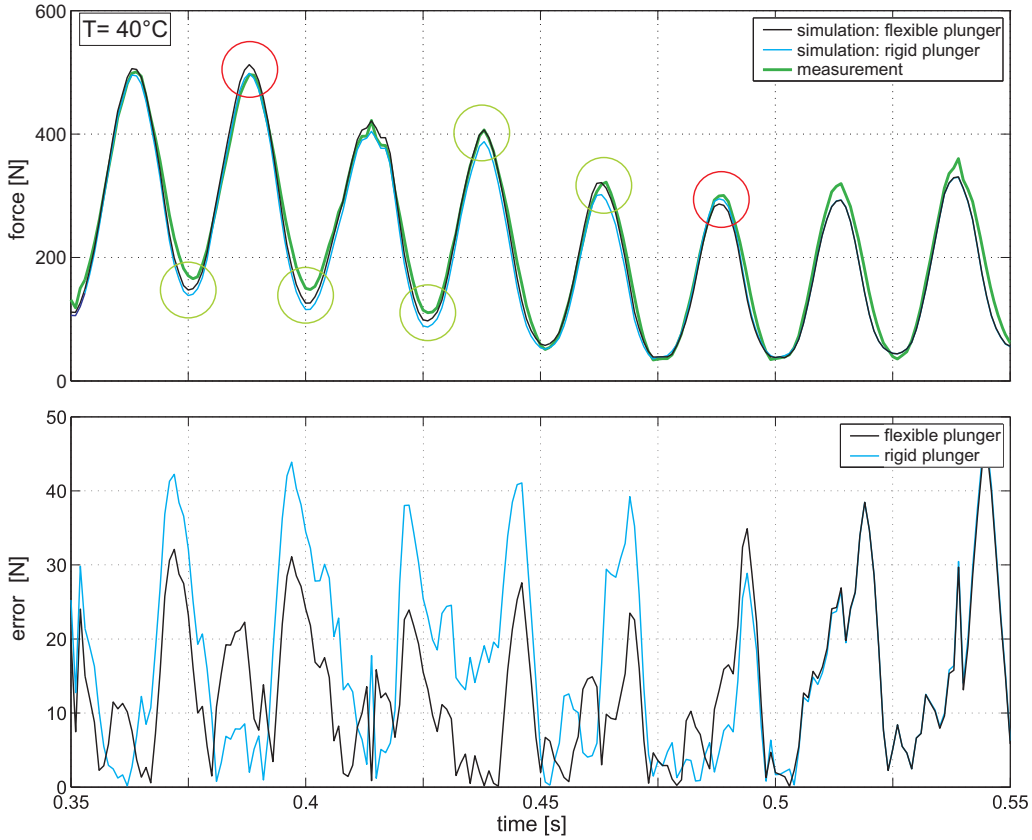


Figure 10.5: Effect of flexible plunger (variable leakage gap)

The double logarithmic interim values W_1 and W_2 are given by:

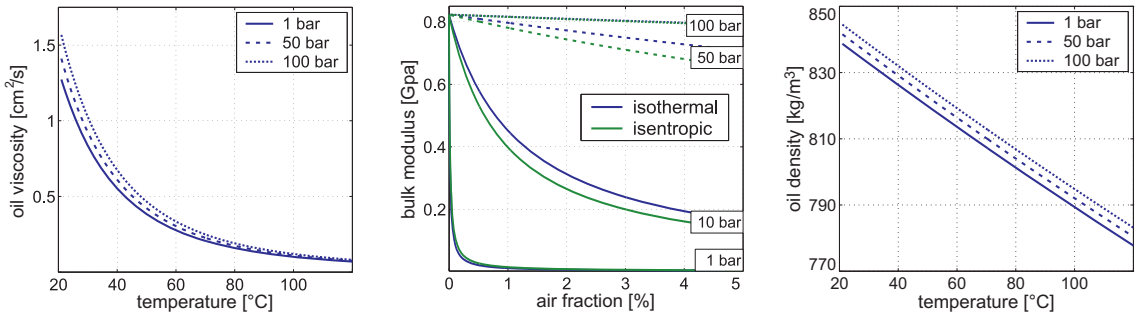
$$W_1 = \log\log(\nu_1\, 10^6 + 0.8)\,, \tag{10.27a}$$
$$W_2 = \log\log(\nu_2\, 10^6 + 0.8)\,. \tag{10.27b}$$

The viscosity ν_x corresponding to T_x is interpolated in two steps:

$$W_x = m\,(\log T_1 - \log T_x) + W_1\,, \tag{10.28a}$$
$$\nu_x = \left(10^{10^{W_x}} - 0.8\right) 10^{-6}\,. \tag{10.28b}$$

Modern motor oils have a lot of additives e.g. detergents, antiwear additives, alkaline additives, and viscosity index improvers. Such non-Newtonian multigraded oils show different viscosity behaviours compared to mineral oils. Nevertheless, the empirical Ubbelohde-Walther equations are applicable for this kind of hydraulic simulation.

Besides viscosity, the bulk modulus is a further physical parameter with a very strong influence on the dynamics of the lash adjuster. As hydraulic fluids typically become aerated in use and aeration has a significant effect on the bulk modulus [173, 191], the air content of the oil has to be considered for the simulation model of the lash adjuster.



(a) Kinematic viscosity (b) Bulk modulus (c) Density

Figure 10.6: Temperature and pressure dependencies of oil properties

Air reduces the bulk modulus of hydraulic fluids because air is much more compressible than oil. Air dissolved in a fluid at high pressure can form bubbles when pressure drops. This phenomenon can cause cavitation. To consider all these effects the compressibility of air and pure oil is combined. For the air volume, an isothermal compression and expansion or an isentropic process can be assumed. With the bulk modulus K_{oil} of pure oil, the heat capacity ratio κ and the volumes $V_{\text{air},0}$ and $V_{\text{oil},0}$ of air and oil at pressure p_0, the bulk modulus for a given air content α and fluid pressure p is given by:

$$K_{\text{fl,isentropic}} = K_{\text{oil}} \frac{1+\alpha}{1+\left(\frac{p_0}{p}\right)^{\frac{1}{\kappa}} \frac{\alpha K_{\text{oil}}}{\kappa p}} \tag{10.29}$$

$$\text{with air fraction} \quad \alpha = \frac{V_{\text{air},0}}{V_{\text{oil},0}},$$

and

$$K_{\text{fl,isothermal}} = K_{\text{oil}} \frac{p}{(1-\tilde{\alpha})\, p + \tilde{\alpha} K_{\text{oil}}} \tag{10.30}$$

$$\text{with} \quad \tilde{\alpha} = \frac{V_{\text{air}}}{V_{\text{air}} + V_{\text{oil}}} = \frac{\alpha}{\alpha + \frac{p}{p_0}}.$$

The pressure and air ratio dependency of the bulk modulus is shown in Figure 10.6(b). High volume fractions of air cause a smooth performance, whereas decreasing air content results in higher plunger forces and stiff behaviour. During engine operation oil and air are permanently mixed. Hence, the air content varies in a low percentage range. For matching experimental results with the simulation this parameter is tuned within a certain range.

To complete the physical properties of the hydraulic fluid, the temperature and pressure dependency of the oil density is considered (cp. Figure 10.6(c)). The temperature dependency is described by the average coefficient α_V of volumetric thermal expansion. The influence of the pressure is measured by the average compressibility $\mathcal{K}$ of the degassed hydraulic fluid. Starting with the density ρ_0 at the reference point

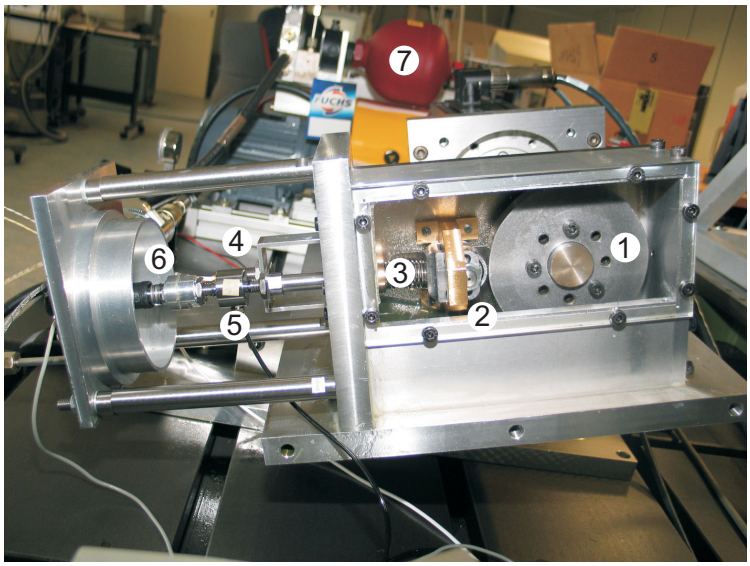
T_0 and p_0, the density as a function of T and p is given by Eq. (10.31):

$$\rho = \rho_0 \frac{1}{1 + \alpha_V (T - T_0)} \frac{1}{1 - \mathcal{K}(p - p_0)} \tag{10.31}$$

In Figure 10.6 the physical properties of the oil in use (Castrol LongLife 2 TopUp 0W-30) are shown.

10.3 Test Rig and Experimental Results

10.3.1 Test Rig



1: cam disc
2: cam roller
3: connection rod
4: displacement sensor
5: force sensor
6: hydraulic lash adjuster
7: oil tank

Figure 10.7: Test rig for the hydraulic lash adjuster

For verification of the simulation model, the dynamics of the lash adjuster are investigated in an experimental setup. The test rig for the investigation of chain tensioners was designed by Engelhardt [50]. It is described by Krüger et al. [124]. With small adaptations it is also possible to use this test rig for hydraulic lash adjusters. The test rig is shown in Figure 10.7. The HLA is inertially fixed and supplied with conditioned oil in a temperature range from 20 to 110° C. The oil feed pump is connected with a heatable tank and the pressure of the pre-conditioned oil can be adjusted by a pressure valve in a range between 0 and 9 bar relative to ambient pressure. In addition, the oil can be heated by a cartridge heater. The dynamic excitation is realised by a stiff cam disc assembly. The circular contour of the cam disc (1) is superposed by a sinusoidal shape generating the desired harmonic translational motion.
The connection rod (3) and the cam roller (2) follow the contour of the cam disc without clearance and excites axially the plunger of the lash adjuster (6). By using different cam discs, different excitation amplitudes can be realised. Contact to the

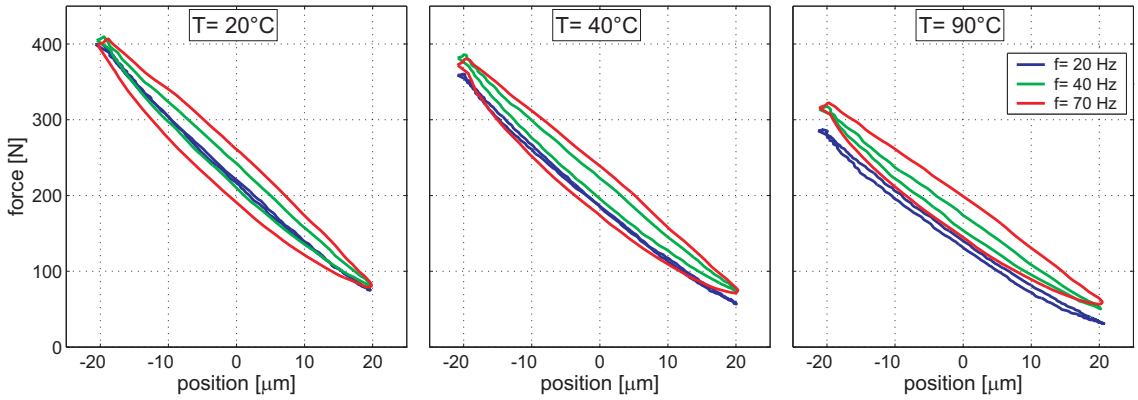


Figure 10.8: Hysteresis curves hydraulic lash adjuster

disc is ensured by a spring. The position of the connection rod is gauged by an eddy current displacement sensor (4), while the contact force between the lash adjuster plunger and the connection rod is measured by a force sensor (5). Furthermore, the pressure in the supply line is recorded.

Hysteresis curves are often used in practice to characterise the performance of non-linear elements such as chain tensioners and lash adjusters. For the present lash adjuster, the plunger force is plotted against the plunger position in Figure 10.8 to illustrate the behaviour in the expansion and compression phase at different temperatures and frequencies. For higher temperatures, the force level is reduced resulting in a less stiff behaviour. The influence of temperature on the damping behaviour is nearly negligible. For higher frequencies, the damping and friction forces become more and more visible.

10.3.2 Boundary Conditions

The choice of the system boundary within the simulation model determines the experimental boundary values. Besides temperature, two boundary conditions are considered: the pressure in the supply line and the kinematic excitation of the system. The measured pressure in the supply line is applied to the hydraulic model as a boundary condition for the low pressure chamber. The kinematic excitation is applied to the body of the force sensor. The force sensor itself is connected to the plunger by a unilateral contact. The stiffness of this connection is chosen according to the force sensor stiffness specified in its data sheet. The plunger force obtained by simulation is compared to the experimental force to verify the simulation model and to adjust unknown parameters.

To use a measured position as a kinematic excitation within the multi-body system simulation in general, the velocity and acceleration have to be calculated. With the help of a Kalman filter [116] the corresponding velocity and acceleration can be derived. Due to measurement errors and signal noise, the Kalman filter provides better accuracy compared to numerical differentiation. The resulting kinematic

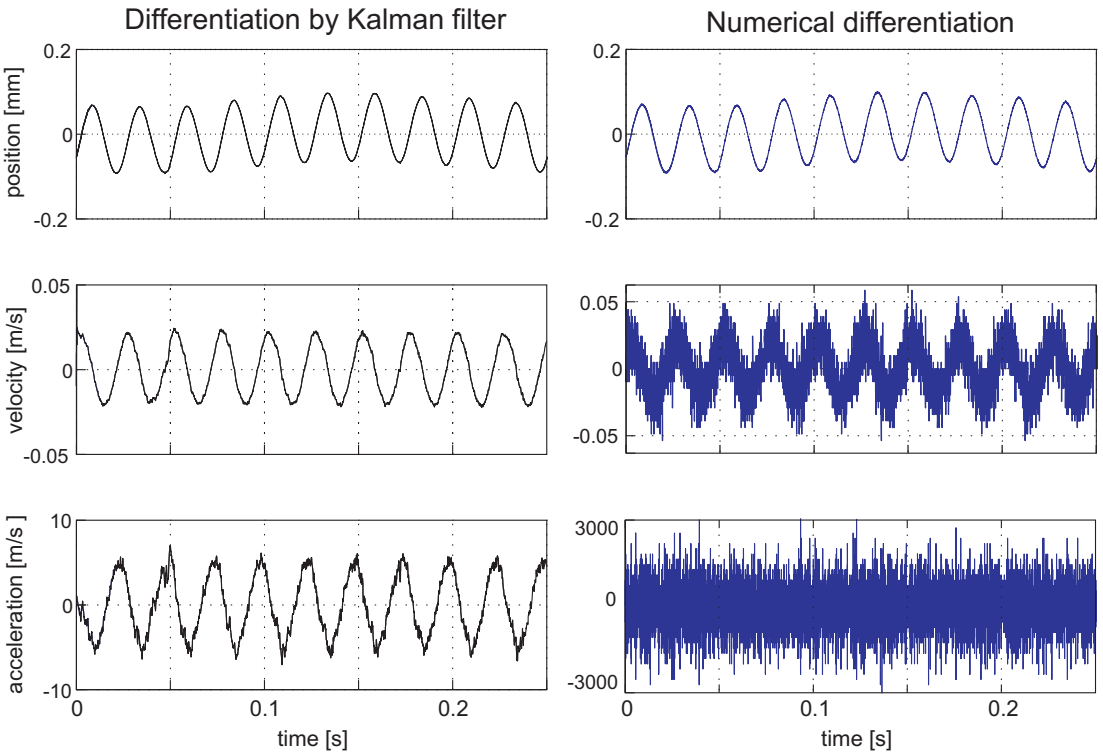


Figure 10.9: Processing measured data: numerical differentiation and Kalman filter

excitation is plotted in Figure 10.9 and compared to the values obtained by numerical differentiation.

10.3.3 Comparison of Experiment and Simulation

Verification of the model and identified parameters is done by the comparing the measured and simulated plunger forces, including several working points at different temperatures, frequencies, and amplitudes. In Figures 10.10 to 10.12 the plunger forces are shown for excitation frequencies from 20 to 70 Hz and 15 μm amplitude. Figure 10.13 presents the results for 50 μm. All simulations are done with the same set of parameters. Only the initial volume of the high pressure chamber is adjusted for every single working point, because one cannot be sure that the starting position of the plunger does not change between two experiments.
There is a remarkable agreement at all working points. The force maxima increase with increasing frequency and decreasing temperature.

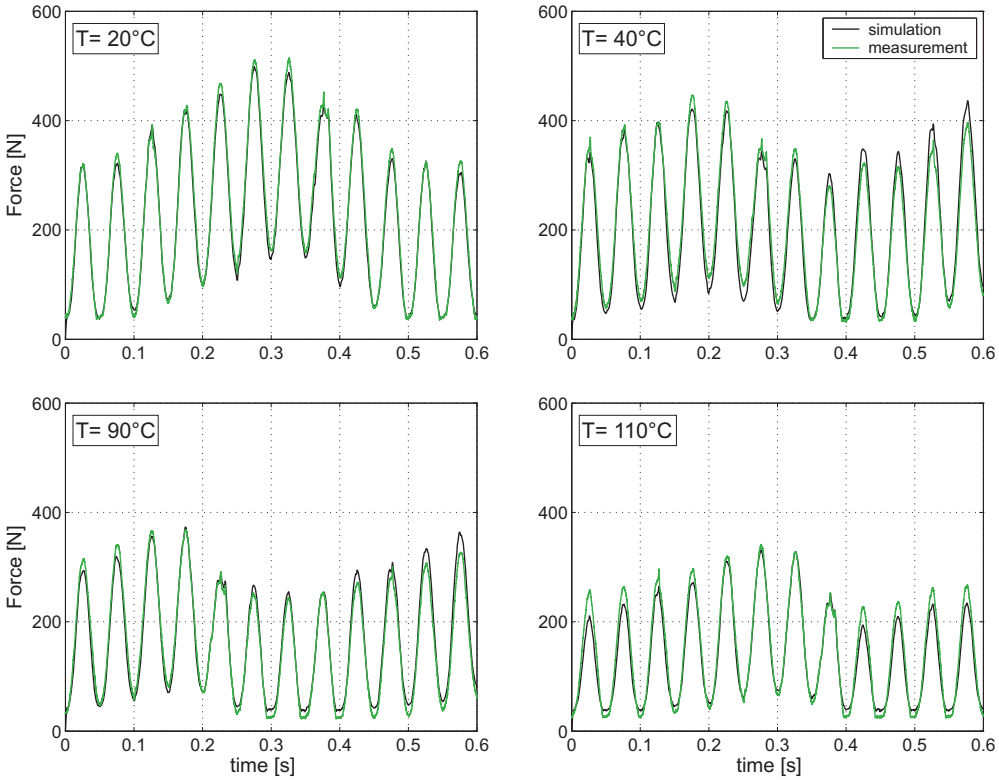


Figure 10.10: Piston amplitude 15 μm and 20 Hz excitation frequency

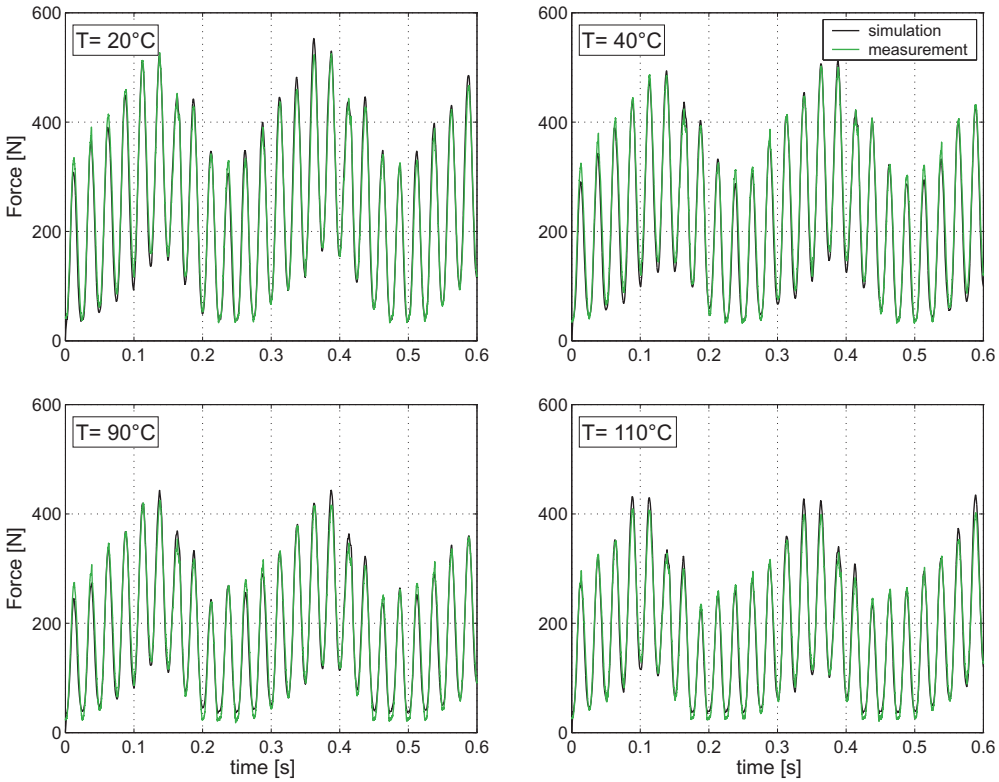


Figure 10.11: Piston amplitude 15 μm and 40 Hz excitation frequency

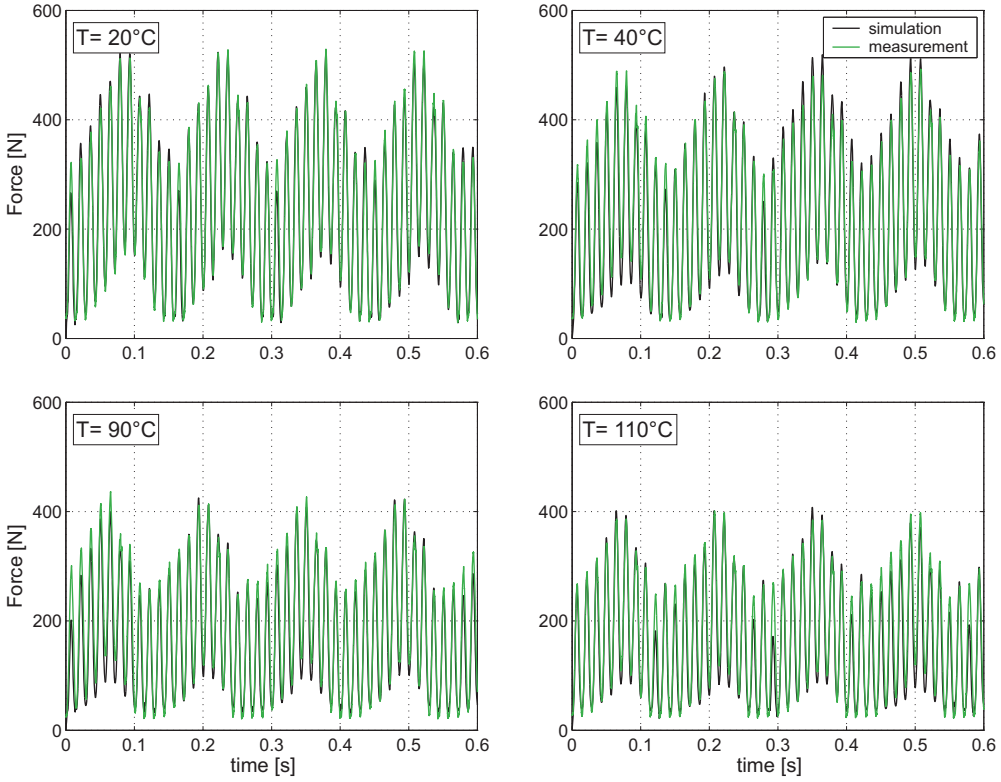


Figure 10.12: Piston amplitude 15 μm and 70 Hz excitation frequency

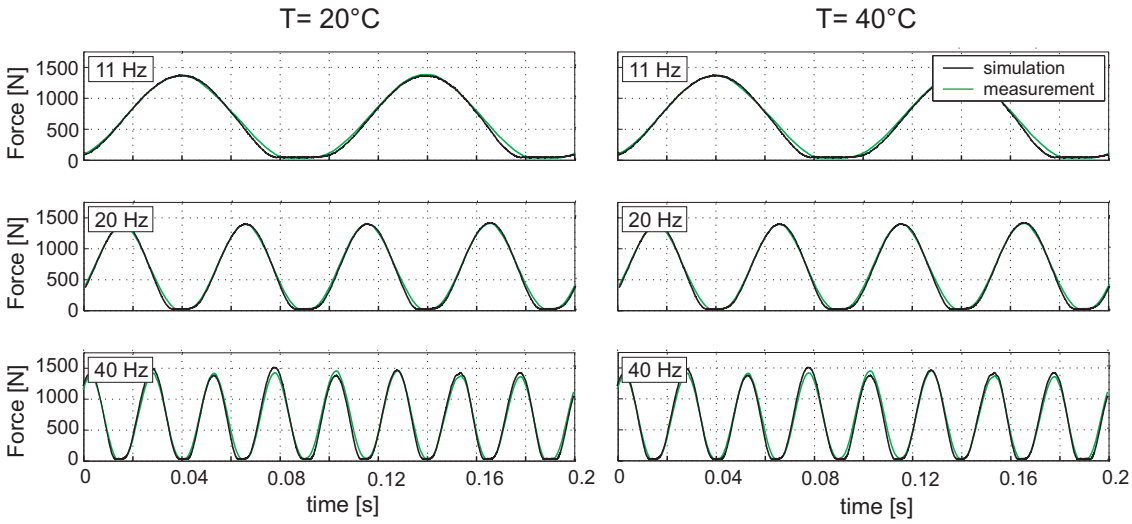


Figure 10.13: Piston amplitude 50 μm

11 AVS Timing Assembly

To analyse the entire AVS valve train, the subsystems presented in the previous sections are merged into one overall simulation model. For studying interactions between the AVS valve train and further subsystems of the timing assembly, the AVS valve train is coupled with a cam phasing system and a chain drive. The entire timing drive is illustrated in Figure 11.1.

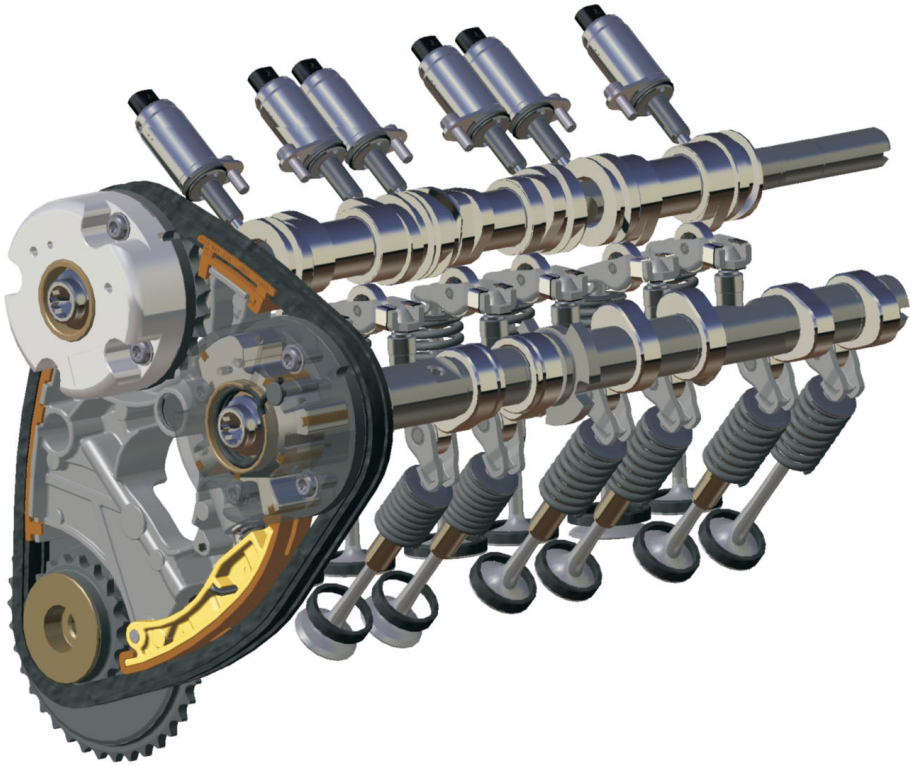


Figure 11.1: Left cylinder bank of AVS valvetrain

11.1 Simulation of the Timing Assembly

11.1.1 Chain Drive and Cam Phaser

Simulation models already exist for the remaining components. Therefore, they are not topics of this study. For further information refer to the following references for comparisons between simulation and measurement for the subsystems.

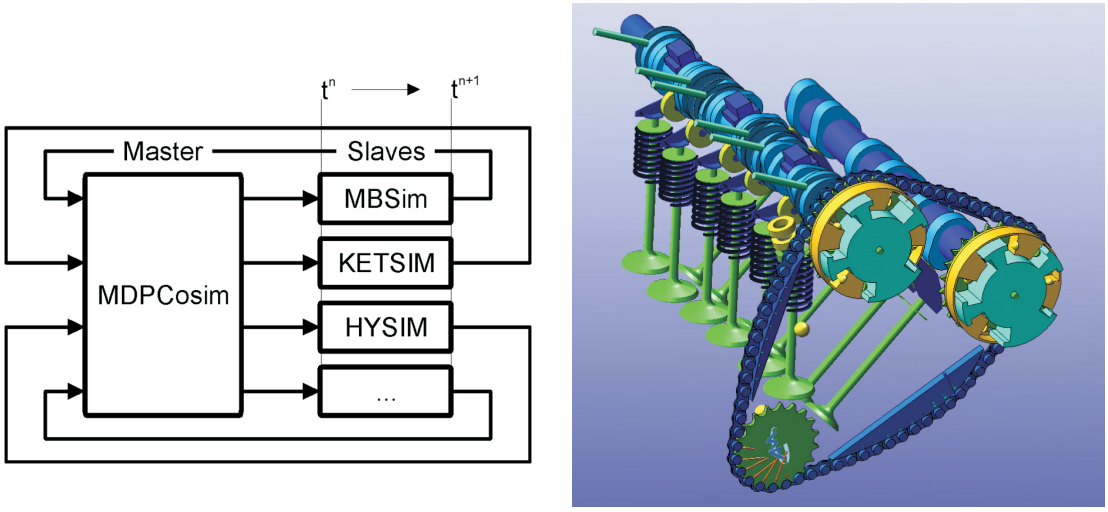
Krüger and Schneider [123, 178, 179] derived a validated simulation model of a hydraulic cam phasing system. In this work a test rig, experimental studies of the dynamical behaviour, simulation models, and simulation results are presented. Simulation models of the cam phaser and the chain tensioner has been adopted from this works. Simulation models and experimental results for different chain tensioners can be found in Krüger et al. [124, 125].
An example for a chain drive can be found in [51]. The theoretical background of chain drive modelling is given by Fritz [70].

11.1.2 Co-Simulation

The conventional approach is to simulate the dynamics of each component e.g. valve train, cam phaser, and chain drive separately from each other mostly using different software tools. The results from one subsystem are used as kinematic or kinetic boundary conditions for the next subsystem. However, in order to analyse the interactions between the subsystems, a combined simulation becomes necessary.

The Institute of Applied Mechanics at the Technische Universität München has developed several highly specialised and efficient computer programs to compute the dynamics of complex multi-body systems including hydraulics [206]. MBSim [177] is used for the AVS valve train and the cam phaser. MBSim is an object orientated software for multi-body systems including hydraulics and control. It differs from standard commercial codes because special attention is devoted to the treatment of uni- and bilateral constraints using non-smooth formulations and methods of convex analysis. MBSim is the basis for current and upcoming research projects including valve trains, continuously variable transmissions, and robotics at the Institute of Applied Mechanics. To calculate the dynamics of the timing chain drive KetSim [70] is used. KetSim is a simulation tool to calculate the dynamics of chain drives. The program is able to simulate bush and roller chains, as well as toothed chains and chain tensioners. The software tool HySim [20] can be used for a more sophisticated modelling of chain tensioners [125].

Each of the simulation programs has been developed to meet different demands of the corresponding subsystems, and therefore it is difficult to include models from other disciplines. One way to solve this problem is to use co-simulation techniques to couple pre-existing simulation tools. The co-simulation method used in this work has been developed by Friedrich [67–69]. The equations of motion of each subsystem are integrated with their respective numerical integrator. The connection between these subsystems is done by communication at discrete macro time steps. During a macro time step the input vector, mainly external forces to the subsystems, is held constant. After integration to the end of the macro time step is done, the subsystems must provide an output vector, e.g. position and velocity of the coupling point. The output vector is used by the master to calculate the new input vector for the subsystems (see Figure 11.2(a)).



(a) Co-Simulation (b) Simulation model of the timing assembly

Figure 11.2: Simulation of the timing assembly

The coupling can be seen as a spring connecting the subsystems. The spring force is calculated by the master. Thereby, the position and velocity information of the end points of the spring are provided by the subsystems or slaves.

The integration of the subsystems is independent of each other. That is the reason why integration of the subsystems can be done parallel on multi-core architectures and why integration time can be speed up. The computational effort of the master is very low compared to the numerical effort for the integration of the subsystems. Thus, the part of serial code is very small and efficient parallelisation is possible. For this work all subsystems are coupled by parallel co-simulation and five threads are used: one for the timing chain drive including the chain tensioner, two for both cam phasers, and two for both AVS valve trains. The overall model is visualised in Figure 11.2(b).

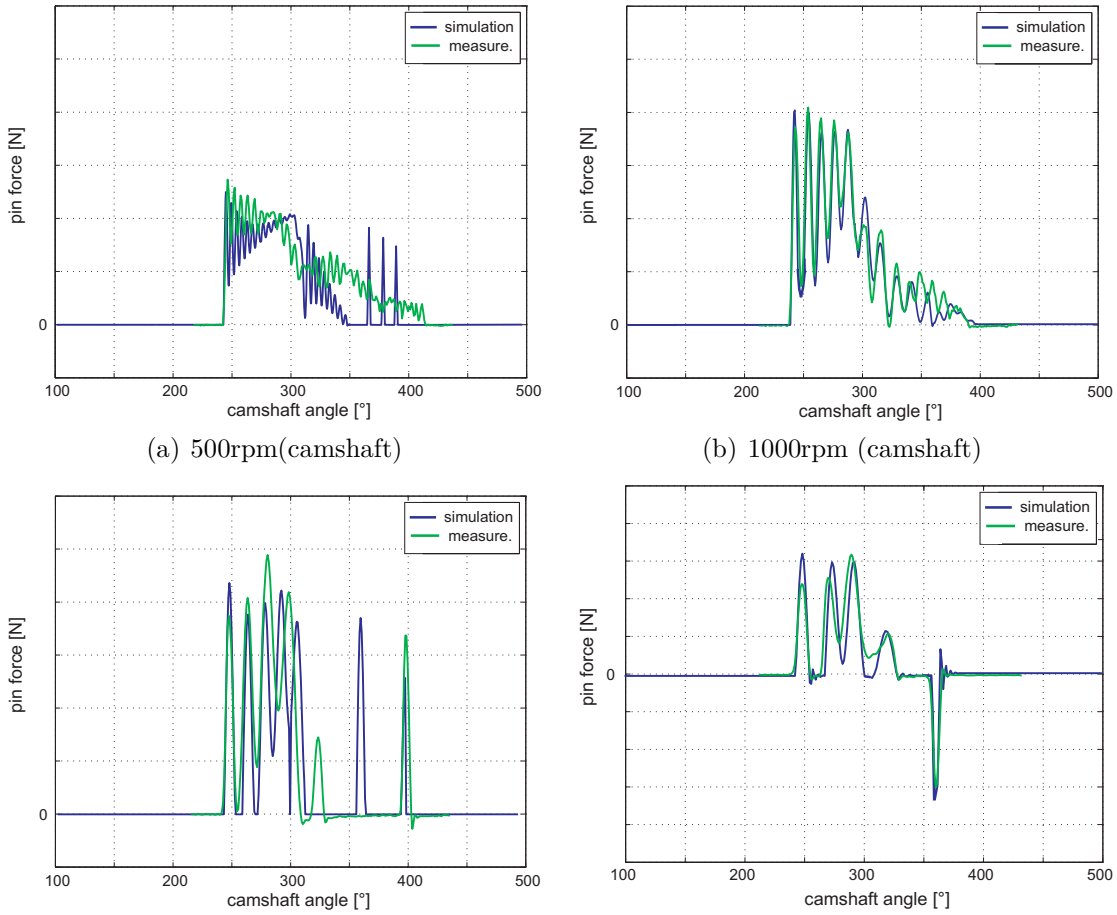
11.2 Results

11.2.1 Pin Forces

To validate the overall AVS valve train simulation model, the forces acting on the pin of the actuator are compared to measurements during a changeover. The pin forces at the lateral contact between pin and groove were measured by Tröster [192] at different speeds using a strain gauge applied to the pin shaft. The results from measurement and simulation are shown in Figure 11.3.

Based on the pin forces, the switching process can be divided into two phases: Phase one is characterised by a permanent contact between pin and groove resulting in an oscillating and continuously positive contact force. This phase is finished by an open contact. Phase two comprises a free-flight phase interrupted by single impacts

between the pin and groove. In general, there is a good correspondence between simulation and experiment in phase one concerning the force level and the frequencies. The force magnitude of single impacts during phase two also matches quite well, whereas the phasing and the transition between both phases differ. Furthermore, the change of sign of single impacts is reproduced correctly by the simulation model. For low speeds contact between the pin and groove always take place at the same side of the groove, whereas for higher speeds (2000rpm) the last impact is located at the opposite groove contour.



(a) 500rpm(camshaft)

(b) 1000rpm (camshaft)

(c) 1500rpm (camshaft)

(d) 2000rpm (camshaft)

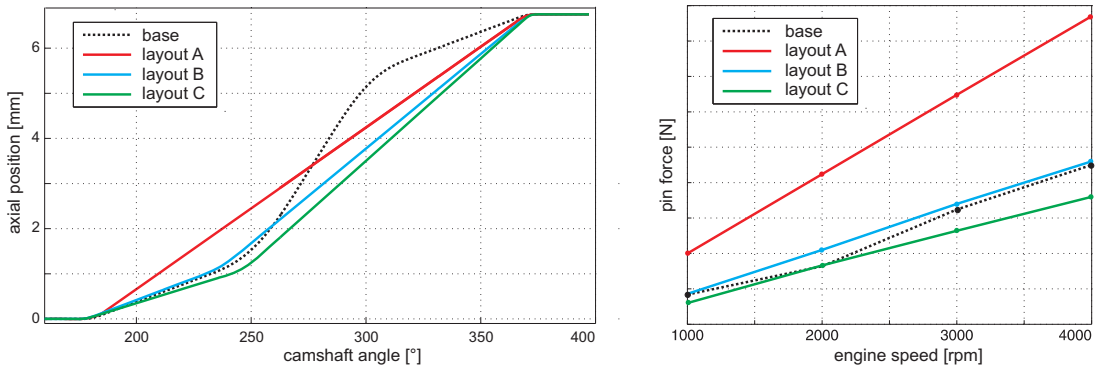
Figure 11.3: Pin force

Having a closer look at the pin force at 500rpm (Figure 11.3(a)), force level and frequency match well during the first part of phase one (up to 300° camshaft angle), even if the simulated frequency is slightly too high. The force amplitudes of the oscillation are significantly too large in the simulation model. This may be one reason that in the simulation model the contact between the pin and groove opens early at 350° camshaft angle. Instead of a permanently closed contact indicated by the experiment, the pin detaches in the simulation and is in a free-flight phase interrupted by three single impacts. Although the force over time differs in measurement and simulation, the exchanged impulse is approximately comparable.
For 1000rpm there is a very good correspondence concerning force level, frequency,

and amplitude as well as the transition between phase one and two. Only at about 350° camshaft angle does the force level differ slightly.
For higher speeds, phase one and two are clearly separated with a distinct free-flight phase. The free-flight phase is terminated by one single impact. In addition, separated impacts occur at the beginning that merges into a permanently closed contact. Both effects can be seen in simulation and experiment. The impact during the free-flight differs, but the ending impact at 350° camshaft angle matches very well. At high speeds (Figure 11.3(d)) the last force peak has a negative sign. This means that at the end of the changeover process the pin hits the opposite side of the groove. This effect is also reproduced well by the simulation model.

11.2.2 Optimisation of the Spiral Groove

The validated simulation model is now used to analyse different contours of the spiral groove in order to reduce the pin forces. In Figure 11.4(a) the developed view of different groove contours are drawn. Three different layouts are under investigation: a contour with S-shape (base), a contour with constant gradient (layout A), and contours with flat initial gradient and steep end gradient (layout B, C).
The initial base contour has been designed with an S-shape. This design is motivated by the claim that at the centre of the S, the pin is "handed over" as smoothly as possible from one side of the groove to the other and pin forces are minimal. Comparing the maximal pin forces obtained by simulation (Figure 11.4(b)), an S-shape seems to not be necessary for minimising the pin forces. Small pin forces are obtained for a flat initial gradient of the spiral groove (layout C).



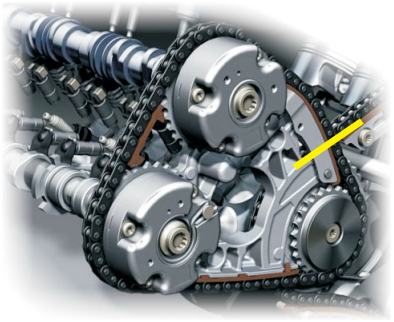
(a) Developed view of spiral groove geometries (b) Lateral Pin force

Figure 11.4: Optimization spiral groove contour

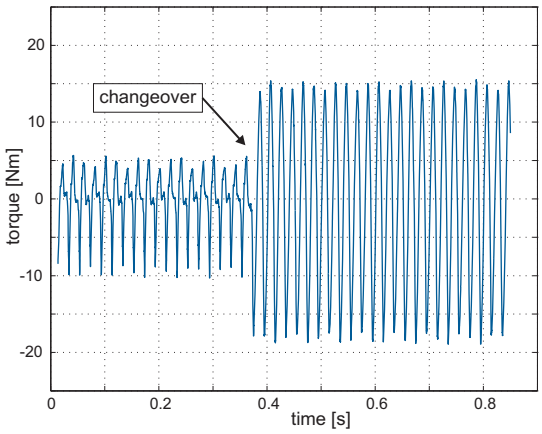
11.2.3 Subsystem Interactions

As an example for subsystem interactions, a changeover process is analysed using the derived simulation model. The effects of a switching process on the chain drive are shown in Figure 11.5. The valve train can be interpreted as a source of excitation

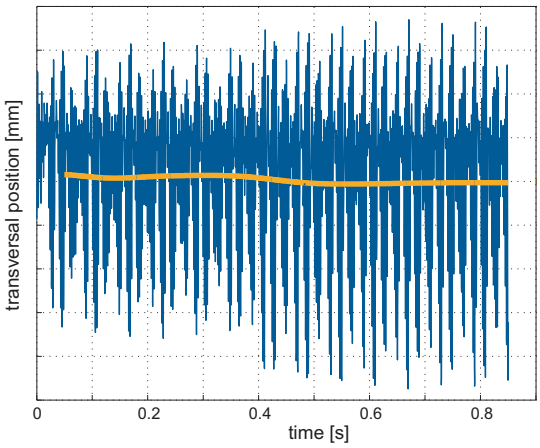
for the chain drive. The cam phasing system acts as a stiff non-linear transfer element and filter. The chain drive with inlet and exhaust cam phasers is shown in Figure 11.5(a). The yellow line indicates the reference position for determining the chain oscillations.



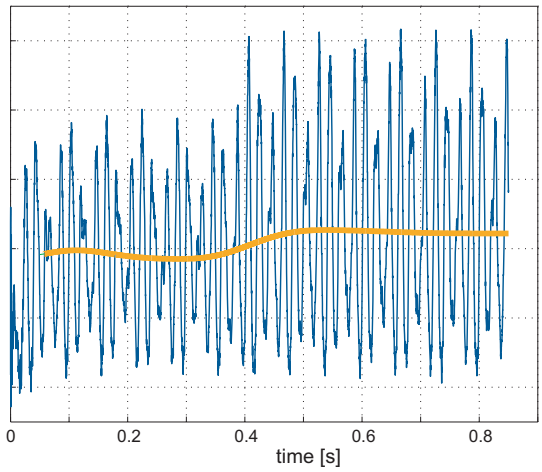
(a) Location of measuring point of the chain forces and oscillations



(b) Torque camshaft



(c) Transversal chain oscillations



(d) Chain forces

Figure 11.5: Interaction of the AVS with the chain drive

The operating cam changes from the partial load configuration with small cam profile to the large profile. Therefore, the camshaft torque increases (Figure 11.5(b)) and the asymmetric excitation due to the asymmetric lift of the two small profiles vanishes. Induced by the changeover and in correspondence with a suddenly rising camshaft torque, the oscillation amplitudes of the transversal chain movement (Figure 11.5(c)) and longitudinal chain force (Figure 11.5(d)) increase by 20 to 25%. No overshoot effect is visible after switching the cam lobes.

As the chain tensioner and the chain drive is designed and laid out for the high force level and excitation with the full load cam profile, the oscillations are normally not critical for the chain drive. However, during partial loads (small cam profile

being active) the force level of the chain tensioner is higher than necessary. A reduction of this chain tensioner force offers an opportunity to reduce friction within the timing assembly, increase the efficiency of the combustion engine, and save CO_2 emission. The simulation model presented in this work is suitable for simulating and evaluating further developments and enhancements of the AVS, such as switching chain tensioners for friction reduction.

Part IV

Conclusion

12 Conclusion

The first part of this work deals with the numerical integration of non-smooth multibody systems (MBS). Non-smooth MBS are characterised by set-valued force laws describing unilateral contacts, friction, and frictional impacts leading to discrete jumps within the system velocities at so called switching points. These points are time instants for which the structure of equations of motion changes. To integrate the underlying measure differential equations of such systems, time-stepping schemes are commonly used. Time-stepping schemes are based on a time discretization of the system dynamics including the constraint conditions. As a consequence, jumps within the system velocities are allowed to fulfil constraints and impact equations.
The fact that time-stepping schemes do not need to resolve switching points makes such schemes ideal for systems with many contacts and switching points. As a drawback, time-stepping schemes are a priori of integration order one and have a constant time discretization. To overcome these disadvantages, a strategy for selecting the step size of a half explicit time-stepping algorithm has been developed within this work. As ODE theory is only applicable during the smooth part of the solution, some adjustments have been made for integration steps containing discontinuities. One of these adjustments is to use the gap velocities to estimate the switching points and use this information to adjust the time discretization. Hence, step size selection can be directed towards robustness and accuracy.
Furthermore, extrapolation methods with time-stepping as the base integration scheme are used to increase the integration order during the smooth part.
Efficiency and robustness of the new integration algorithm with step size selection and higher integration order for non-smooth systems developed within his work, has been tested using various examples with academic (e.g. woodpecker toy, many balls in a barrel) and industrial backgrounds (e.g. variable valve trains, pushbelt CVT). Extrapolation to orders higher than two or three has been shown to be inefficient for several technical systems with weak discontinuities (e.g. due to discrete controllers or imported contours from CAD), even in the smooth part.
As extrapolation methods themselves contain algorithmic parallelism, parallel computing techniques are used to speed up integration by computing all sub-step series simultaneously. The achieved accelerations due to parallelisation are very close to the maximal theoretical values.

The second part of this work presents the modelling and dynamical analysis of a variable valve train, namely the Audi valve lift system (AVS) using the methods presented in the first part of this work. Moreover, modelling approaches are presented for valve trains with a focus on contact mechanics—for example the contact between cam and roller.

In a first step, dynamic models of the components of the AVS valve train are derived and validated using specifically designed test rigs. The AVS specific switching unit consists of a cam element fixed for co-rotation with the camshaft by an involute gearing and an electromagnetic actuator. Besides dynamical investigations, special friction measurements for the gearing have been carried out. For the electromagnetic actuator a simulation model based on an analytical solution of Maxwell's field theory equations has been derived and validated with experimental data. Moreover, a new model of a mass carrying coil spring based on a curved beam including coil to coil contacts has been derived and validated for the simulation of valve springs. In addition, modelling and validation of the hydraulic lash adjuster is shown.
In the next step, the validated simulation models of the components are integrated into an overall simulation model of the entire AVS valve train. This model is used to predict and optimise mechanical forces on the AVS components during a changeover. Finally, the dynamics of the entire timing assembly consisting of chain drive, cam phasing system, and AVS valve train are analysed using parallel co-simulation techniques while concentrating on the interaction between the components during a switching process.

Bibliography

[1] Special Issue on OpenMP. *The International Journal of Parallel Programming*, 36(3):289–346, 2008.

[2] Milton Abramowitz and Irene Stegun. *Handbook of Mathematical Functions*. Dover Publications, 1972.

[3] Vincent Acary and Bernard Brogliato. *Numerical Methods for Nonsmooth Dynamical Systems: Applications in Mechanics and Electronics*, volume 35 of *Lecture Notes in Applied and Computational Mechanics*. Springer Verlag, 2008.

[4] Kazuhiro Akima, Kazuyuki Seko, Wataru Taga, Kenji Torii, and Satoshi Nakamura. Development of new low fuel consumption 1.8l i-VTEC gasoline engine with delayed intake valve closing. *SAE Technical Paper 2006-01-0192*, 2006.

[5] P. Alart and A. Curnier. A mixed formulation for frictional contact problems prone to Newton like solution methods. *Computer Methods in Applied Mechanics and Engineering*, 92:353–375, 1991.

[6] Pierre Alart, Oliver Maisonneuve, and R. Tyrrell Rockafellar. *Nonsmooth Mechanics and Analysis*, volume 12 of *Advances in Mechanics and Mathematics*. Springer Verlag, 2006.

[7] William Conrad Albertson and Frederick J. Rozario. Cylinder deactivation apparatus, 2003. Patent US 6557518.

[8] Albrecht Hußmann. *Schwingungen in schraubenförmigen Ventilfedern*. PhD thesis, Technische Hochschule Berlin, 1938.

[9] Johannes Altenbach and Holm Altenbach. *Einführung in die Kontinuumsmechanik*. B. G. Teubner, 1994.

[10] Martin Arnold. *Zur Theorie und zur Lösung von Anfangswertproblemen für differentiell-algebraische Systeme von höherem Index*. Number 264 in Fortschritt-Berichte VDI, Reihe 20. VDI Verlag, 1998.

[11] G. Backhaus. *Deformationsgesetze*. Akademie-Verlag, 1982.

[12] L. F. Bakaj. *Rechnerische Simulation des dynamischen Verhaltens eines OHC-Schlepphebelventiltriebs mit hydraulischem Ventilspielausgleich*. Number 112 in Fortschritt-Berichte VDI, Reihe 12. VDI Verlag, 1988.

[13] Cristian Barbarosie. Representation of divergence-free vector fields. *Quarterly of Applied Mathematics*, 69:309–316, 2011.

[14] Richard Van Basshuysen and Fred Schafer. *Internal Combustion Engine Handbook: Basics, Components, Systems, and Perspectives*. SAE International, 2004.

[15] J. Baumgarte. Stabilization of constraints and integrals of motion in dynamical systems. *Computer Methods in Applied Mechanics and Engineering*, 1:1–16, 1972.

[16] Ernst Becker and Wolfgang Bürger. *Kontinuumsmechanik*. B. G. Teubner, 1975.

[17] Anouar Belahcen. *Magnetoelasticity, magnetic forces and magnetostriction in electrical machines*. PhD thesis, Helsinki University of Technology, 2004.

[18] Wolf-Dieter Bensinger. *Die Steuerung des Gaswechsels in schnellaufenden Verbrennungsmotoren*. Springer Verlag, 1968.

[19] Lucio Bernard, Andrea Ferrari, Damiano Micelli, Aldo Perotto, Rinaldo Rinolfi, and Francesco Vattaneo. Electro-hydraulic valve control with multiair technology. *MTZ worldwide Edition*, 70(12):892–899, 2009.

[20] Fredrik Borchsenius. *Simulation Ölhydraulischer Systeme*. Fortschrittberichte VDI, Reihe 8. VDI Verlag, 2003.

[21] Fredrik Borchsenius and Friedrich Pfeiffer. New hydraulic system modelling. *Journal of Vibration and Control*, 10(10):1493–1515, 2004.

[22] Carlo L. Bottasso, Daniel Dopico, and Lorenzo Trainelli. On the optimal scaling of index three daes in multibody dynamics. *Multibody System Dynamics*, 19(1-2):3–20, 2008.

[23] Dietrich Braess. *Finite Elemente*. Springer Verlag, 2003.

[24] C. Brands. *Nockenwellendynamik: Einfluss der konstruktiven Steifigkeit von Nockenwelle und Lagerung sowie der nichtlinearen hydrodynamischen Schmierfilmsteifigkeit und Dämpfung auf die dynamischen Eigenschaften des Nockenwellenbetriebes*, volume 709 of *Vorhaben Nr. 693, Abschlussberichte*. Forschungsvereinigung Verbrennungskraftmaschinen, 2001.

[25] Bernard Brogliato. *Nonsmooth Mechanics*. Springer Verlag, 1999.

[26] Bernard Brogliato, A.A. ten Dam, Laetitia Paoli, Frank Ge´not, and M. Abadie. Numerical simulation of finite dimensional multibody nonsmooth mechanical systems. *ASME Applied Mechanics Reviews*, 55(2):107–150, 2002.

[27] Ilja N. Bronstein, K. A. Semendjajew, Gerhard Musiol, and Heiner Muehlig. *Taschenbuch der Mathematik*. Harri Deutsch, 2008.

[28] Claus Brüstle and Dietmar Schwarzenthal. Variocam plus - a highlight of the porsche 911 turbo engine. *SAE Technical Paper 2001-01-0245*, 2001.

[29] Manuel Calvo, Juan Ignacio Montijano, and Luis Rández. On the solution of discontinuous ivps by adaptive runge-kutta codes. *Numerical Algorithms*, 33(1-4):163–182, 2003.

[30] Jeffrey Cash. Review paper: Efficient numerical methods for the solution of stiff initial-value problems and differential algebraic equations. *Proceedings of the Royal Society A*, 459:797–815, 2003.

[31] Alberto Castigliano. *Théorie de l'équilibre des systèmes élastiques et ses applications*. Negro, Turin, 1880.

[32] Chingyao Chan and Albert P. Pisano. Dynamic model of a fluctuating rocker-arm ratio cam system. *Journal of Mechanisms Transmissions and Automation in Design*, 109(3):356–365, 1987.

[33] B. Chapman, G. Jost, and R. van der Pas. *Using OpenMP - Portable Shared Memory Parallel Programming*. MIT Press, October 2007.

[34] Jan Clauberg. Implementierung und Abgleich eines effizienten kontinuierlichen Federmodells mit nichtglatter Kontaktmodellierung für die Ventiltriebssimulation. Master's thesis, TU München, 2009.

[35] C. Degenhardt. Nockentrieb Reibungsverhältnisse. *FVV Forschungsberichte, Heft 483-1*, 1991.

[36] Wolfgang Demtröder. *Experimentalphysik 2: Elektrizität und Optik*. Springer Verlag, 2008.

[37] Dirk Denger and Karsten Mischker. The electro-hydraulice vavetrain system. *MTZ worldwide Edition*, 65(12):978–987, 2004.

[38] Stefan Dengler. *Konzeption, Ausarbeitung und Bewertung neuer Lösungsansätze für mechanisch variable Ventiltriebe*. PhD thesis, Universität Karlsruhe, 2004.

[39] P. Deuflhard and F. Bornemann. *Numerische Mathematik 2: Gewöhnliche Differentialgleichungen*. Gruyter, 2008.

[40] P. Deuflhard and A. Hohmann. *Numerical Analysis in Modern Scientific Computing: An Introduction. Second edition.* Texts in Applied Mathematics 43. Springer Verlag, 2003.

[41] Deutsches Insitut für Normung. Prüfung von Mineralölen und verwandten Stoffen - Bestimmung des Viskosität-Temperatur-Verhaltens. Technical report, DIN 51563 : 2011-04, 2011.

[42] Jochen Dilthey. *Möglichkeiten zur Verbesserung der drosselfreien Laststeuerung beim ottomotorischen Prozess*. PhD thesis, RWTH Aachen, 2004.

[43] Mathias Roman Dreyer. *Untersuchung zur nichtlinearen Mechanik von Schraubendruckfedern*. Der Andere Verlag, Osnabrück, 2004.

[44] Markus Duesmann. Innovative valve train systems. *Spectrum - Technology Highlights and R&D Activities at FEV*, 19:1–3, 2002. Official newsletter of FEV.

[45] Markus Duesmann and Wolfgang Salber. Der Weg zur drosselfreien Laststeuerung: Vom Ventilabschalter bis zum elektromechanischen Ventiltrieb. In Stefan Pischinger, editor, *Variable Ventilsteuerung, Band 1*, pages 32–46. expert-Verlag, 2002.

[46] Bruno Eck. *Technische Strömungslehre*. Springer Verlag, 1966.

[47] Günther Eggerath, Uwe Geiger, Michael Haas, and Lothar Schimonsky. Schaltbare Ventiltriebskomponenten. *Motortechnische Zeitschrift*, 66(7/8):558–565, 2005.

[48] Edda Eich-Soellner and Claus Führer. *Numerical Methods in Multibody Dynamics*. B. G. Teubner, 1998.

[49] Axel Eiser, Joachim Doerr, Michael Jung, and Stephan Adam. Der neue 1,8-L-TFSI-Motor von Audi - Teil 1: Grundmotor und Thermomanagemen. *Motortechnische Zeitschrift*, 72:466–474, 2011.

[50] Thomas Engelhardt. *Dynamik von Steuer- und Ventiltrieben*. Number 656 in Fortschritt-Berichte VDI, Reihe 12. VDI Verlag, 2007.

[51] Thomas Engelhardt, Andreas Hösl, Wolfram Lebrecht, Heinz Ulbrich, and Friedrich Pfeiffer. Simulation of timing chain drives using KetSim. In *ASME Internal Combustion Engine Division Technical Conference, Long Beach Bd. ICEF2004-877*, 2004.

[52] H.-D. Erdmann, Th. Heiduk, P.A. Klumpp, and G. Mendl. An adaptive dynamic model for the virtual engine. *ATZ autotechnology*, 4, 2001.

[53] Alexander Ewald. Experimentelle Untersuchungen von Reibkontakten an einer Nockenwelle eines variablen Ventiltriebs. Master's thesis, TU München, 2008.

[54] Joachim Fetzer, Martin Haas, and Stefan Kurz. *Numerische Berechnung elektromagnetischer Felder*. expert-Verlag, 2002.

[55] Dietmar Findeisen and Franz Findeisen. *Ölhydraulik - Handbuch für die hydrostatische Leistungsübertragung in der Fluidtechnik*. Springer Verlag, 1994.

[56] Torsten Fließbach. *Lehrbuch zur Theoretischen Physik 2. Elektrodynamik*. Spektrum Akademischer Verlag, Heidelberg, 5. edition, 2008.

[57] Torsten Fließbach. *Lehrbuch zur Theoretischen Physik 1. Mechanik*. Spektrum Akademischer Verlag, Heidelberg, 6. edition, 2009.

[58] Rudolf Flierl, Reinhard Hofmann, Christian Landerl, Theo Melcher, and Helmut Steyer. Der neue BMW Vierzylinder-Ottomotor mit VALVETRONIC Teil 1: Konzept und konstruktiver Aufbau. *Motortechnische Zeitschrift*, 62(6):450–463, 2001.

[59] Rudolf Flierl and Manfred Klüting. The third generation of valvetrains - new fully variable valvetrains for throttle-free load control. *SAE Technical Paper 2000-01-1227*, 2000.

[60] Rudolf Flierl, Stephan Schmitt, Din Wabbals, and Gerd Kleinert. Voll variabler Ventiltrieb und variable Verdichtung. *Motortechnische Zeitschrift*, 70:596–602, 2009.

[61] L. J. Flynn. Intel halts development of 2 new microprocessors. *The New York Times*, May 8 2004.

[62] M. Förg, L. Neumann, Th. Geier, and H. Ulbrich. r-factor Strategies for the Augmented Lagrangian Approach in Multi-Body Contact Mechanics. In *Proceedings of the III European Conference on Computational Mechanics*, Lisbon, Portugal, 2006.

[63] M. Förg, R. Zander, and H. Ulbrich. A Framework for the Efficient Simulation of Spatial Contact Problems. In *Proceedings of the ECCOMAS Conference on Multi-Body Systems*, Milano, Italy, 2007.

[64] Martin Förg. *Mehrkörpersysteme mit mengenwertigen Kraftgesetzen – Theorie und Numerik*. PhD thesis, TU München, Düsseldorf, 2007.

[65] Martin Förg, Thomas Engelhardt, and Heinz Ulbrich. Comparison of differnt contact models within valve train simulations. In *Proceedings of ACMD06*, 2006.

[66] W. C. Franke. Progress in simulation of valve train dynamics. In *Proceedings of the 1956 11th ACM national meeting, New York*, pages 87–90, 1956.

[67] Markus Friedrich. *Parallel Co-Simulation for Mechatronic Systems*. PhD thesis, Technische Universität München, 2011.

[68] Markus Friedrich, Markus Schneider, and Heinz Ulbrich. A parallel co-simulation for mechatronic systems. In *1st Joint International Conference on Multibody System Dynamics, Lappeenranta, Finland*, 2010.

[69] Markus Friedrich and Heinz Ulbrich. A parallel co-simulation for multi body systems. In *2nd South-East European Conference on Computational Mechanics, Rhodes, Greece*, 2009.

[70] Peter Fritz. *Dynamik schnellaufender Kettentriebe*. Number 253 in VDI Fortschritt-Berichte. VDI Verlag, 1997.

[71] C.-P. Fritzen and Mathias Roman Dreyer. Numerical and experimental investigations of the dynamics of valve trains in combustion engines. In *Proceedings ISMA 23*, Noise and Vibration Engeneering, pages 259–267, Leuven, 1998.

[72] Anton Fritzer. *Nichtlineare Dynamik von Steuertrieben*. Number 176 in VDI Fortschritt-Berichte. VDI Verlag, 1992.

[73] K. Funk and F. Pfeiffer. A Time-Stepping Algorithm for Stiff Mechanical Systems with Unilateral Constraints. In *Proc. of the International Conference on Nonsmooth/Nonconvex Mechanics with Applications in Engineering*, pages 307–314, Thessaloniki, Greece, 2002.

[74] Kilian Funk. *Simulation eindimensionaler Kontinua mit Unstetigkeiten*. Number Nr. 294 in Fortschrittberichte VDI, Reihe 18. VDI Verlag, Düsseldorf, 2004.

[75] Charles William Gear, Ben Leimkuhler, and Gopal K. Gupta. Automatic integration of euler-lagrange equations with constraints. *Journal of Computational and Applied Mathematics*, 12-13:77–90, 1985.

[76] Christop Glocker. *Dynamik von Starrkörpersystemen mit Reibung und Stößen*. VDI Verlag, 1995.

[77] Christop Glocker. Models of non-smooth switches in electrical systems. *International Journal of Circuit Theory and Applications*, 33(3):205–234, 2005.

[78] O. Göhner. Schubspannungsverteilung im Querschnitt einer Schraubenfeder. *Archive of Applied Mechanics*, 1(5):619–644, 1930.

[79] O. Göhner. Schubspannungsverteilung im Querschnitt eines gedrillten Ringstabs mit Anwendung auf Schraubenfedern. *Archive of Applied Mechanics*, 2(1):1–19, 1931.

[80] O. Göhner. Die Berechnung zylindrischer Schraubenfedern. *Zeitschrift des VDI*, 76(11):269, 1932.

[81] Donald T. Greenwood. *Advanced Dynamics*. Cambridge University Press, 2003.

[82] Michael Grigo, Rainer Wurms, Ralf Budack, Jan Helbig, Zsolt Langa, and Wolfgang Trost. Der neue 2,0-l-TFSI-Motor mit Audi valvelift system. *Motortechnische Zeitschrift*, pages 30–34, 2008.

[83] Karl-Heinrich Grote and Jörg Feldhusen. *Dubbel - Taschenbuch für den Maschinenbau*. Springer Verlag, 2001.

[84] E. Hairer, Ch. Lubich, and M. Roche. The numerical solution of differential-algebraic systems by runge-kutta methods. In *Lecture Notes in Math. 1409*. Springer Verlag, 1989.

[85] E. Hairer, S.P. Norsett, and G. Wanner. *Solving ordinary differential equations. I, Second Edition*. Springer Series in Computational Mathematics. Springer Verlag, 2003.

[86] Ernst Hairer, Christian Lubich, and Michel Roche. *The Numerical Solution of Differential-Algebraic Systems by Runge-Kutta Methods*, volume 1409 of *Lecture Notes in Mathematics*. Springer Verlag, 1989.

[87] Ernst Hairer and Gerhard Wanner. *Solving Ordinary Differential Equations II: Stiff and Differential-Algebraic Problems*, volume 14 of *Springer Series in Computational Mathematics*. Springer Verlag, 1991.

[88] Georg Hamel. *Theoretische Mechanik*. Springer Verlag, 1949.

[89] Shervin Hanachi and Ferdinand Freudenstein. The development of a predictive model for the optimization of high-speed cam-follower systems with coulomb damping internal friction and elastic and fluidic elements. *Journal of Mechanisms Transmissions and Automation in Design*, 108:506–515, 1986.

[90] Wilhelm Hannibal. *Vergleichende Untersuchung verschiedener variabler Ventilsteuerungen für Serien-Ottomotoren*. PhD thesis, Universität Stuttgart, 1993.

[91] Wilhelm Hannibal, Rudolf Flierl, and Lutz Stiegler. Overview of current continuously variable valve lift systems for four-stroke spark-ignition engines and the criteria for their design ratings. *SAE Technical Paper 2004-01-1263*, 2004.

[92] Kiyoshi Hatano, Kazumasa Iida, Hirohumi Higashi, and Shinichi Murata. Development of a new multi-mode variable valve timing engine. *SAE Technical Paper 930878*, 1993.

[93] C. T. Hatch and A. P. Pisano. Modeling, simulation, and modal analysis of a hydraulic valve lifter with oil compressibility effects. *Journal of Mechanical Design*, 113(1):46–54, 1991.

[94] Klaus Häufle. *Untersuchungen des Betriebsverhaltens von Ventilfedern mit Hilfe eines transienten FEM-Simulationsmodells für Ventilsteuerungen*, volume 218 of *Fortschritt-Berichte VDI*. VDI Verlag, 1995.

[95] Thomas Heiduk. *Methoden zur Analyse von Schwingungen und dynamischen Kräften in Steuerungsantrieben von Verbrennungsmotoren für PKW*. PhD thesis, RWTH Aachen, 1995.

[96] Thomas Heiduk, Alex Eiser, Michael Fitzen, Wolfgang Hatz, and Johann Mendle. Der neue V6-2,8-l-FSI mit Audi-Valvelift-System. In *15. Aachener Kolloquium - Fahrzeug- und Motorentechnik*, 2006.

[97] Thomas Heiduk, Michael Kuhn, Maximilian Stichlmeir, and Florian Unselt. Der neue 1,8-L-TFSI-Motor von Audi - Teil 2: Gemischbildung, Brennverfahren und Aufladung. *Motortechnische Zeitschrift*, 72:596–604, 2011.

[98] Conrad P. Heins. *Bending and torsional design in structural members*. Lexington Books, 1975.

[99] Marcus Herrmann. *Entwurf, Berechnung und Regelung magnetischer Reluktanzaktoren*. Ingenieurwissenschaften. Verlag Dr. Hut, 2008.

[100] Heinrich Hertz. Über die Berührung fester elastischer Körper. *Journal für reine und angewandte Mathematik*, 92:156–171, 1881.

[101] Vic Hillier. *Hillier's Fundamentals of Motor Vehicle Technology*. Nelson Thornes, 6 edition, 2004.

[102] P. Hofbauer. *Nockenoptimierung für den Ventiltrieb von Fahrzeugmotoren*. PhD thesis, RWTH Aachen, 1981.

[103] Hermann Hoffmann, Adam Loch, Richard Widmann, Gerhard Kreusen, Daniel Meehsen, and Martin Rebbert. Zylinderabschaltung für Ventiltriebe mit Rollenschlepphebeln. *Motortechnische Zeitschrift*, 70:302–307, 2009.

[104] Gerhard Holzapfel. *Nonlinear Solid Mechanics: A Continuum Approach for Engineering*. John Wiley & Sons Ltd., 2005.

[105] Takefumi Hosaka and Minoru Hamazaki. Development of the variable valve timing and lift (VTEC) engine for the honda NSX. *SAE Technical Paper 910008*, 1991.

[106] Andreas Hösl. *Dynamiksimulation von Steuerkettentrieben*. Number 618 in VDI Fortschritt-Berichte. VDI Verlag, 2005.

[107] Robert Huber. Dynamic model of a mass carrying spring based on fundamental continuum mechanics. Master's thesis, TU München, 2011.

[108] Robert Huber, Peter Klumpp, and Heinz Ulbrich. Dynamic analysis of the audi valvelift system. *SAE International Journal of Engines*, 3(1):839–849, 2010.

[109] Robert Huber and Heinz Ulbrich. Simulation of a valve train using non-smooth mechanics. *SAE International Journal of Engines*, 1(1):208–217, 2009.

[110] Helmut Hütten. *Schnelle Motoren seziert und frisiert*. Motorbuch-Verlag, 1994.

[111] I. E. Idelchik. *Handbook of Hydraulic Resistance*. Jaico Publishing House, 3 edition, 2005.

[112] Hartmut Janocha. *Actuators: Basics and Applications*. Springer Verlag, 2004.

[113] Daniel J. Jendritza. *Technischer Einsatz neuer Aktoren*. expert-Verlag, Renningen-Malmsheim, 1995.

[114] K.L. Johnson. *Contact Mechanics*. Cambridge University Press, 1987.

[115] Eberhard Kallenbach, Rüdiger Eick, Peer Quendt, Tom Ströhla, Karsten Feindt, and Matthias Kallenbach. *Elektromagnete - Grundlagen, Berechnung, Entwurf und Anwendung*. B. G. Teubner, 2 edition, 2003.

[116] Rudolf E. Kalman. A new approach to linear filtering and prediction problems. *Transaction of the ASME, Journal of Basic Engineering*, pages 35–45, 1960.

[117] H.-J. Kalser, R. Deges, D. Schwarz, and J. Meyer. Investigations on valve train noise in multi-valve engines. *SAE Technical Paper 911062*, 1991.

[118] Kurt Kirsten. The engine - understanding in its entirety. Technical Report 14, Schaeffler SYMPOSIUM, 2010.

[119] Carsten Kopp. *Variable Ventilsteuerung für Pkw-Dieselmotoren mit Direkteinspritzung*. PhD thesis, Otto-von-Guericke-Universität Magdeburg, 2006.

[120] Arnulf Kost. *Numerische Methoden in der Berechnung elektromagnetischer Felder*. Springer Verlag, 1994.

[121] P. Kreuter and G. Maas. Influence of hydraulic valve lash adjusters on the dynamic behavior of valve trains. *SAE Technical Paper 870086*, 1987.

[122] P. Kreuter and F. Pischinger. Valve train calculation model with regard to oil film effects. *SAE Technical Paper 850399*, 1985.

[123] Karin Krüger. *Einfluss hydraulischer Nockenwellenversteller auf die Dynamik von Ventil- und Steuertrieben*. Verlag Dr. Hut, 2009.

[124] Karin Krüger, Thomas Engelhardt, Lucas Ginzinger, and Heinz Ulbrich. Dynamical analysis of hydraulic chain tensioners - experiment and simulation. *SAE Technical Paper 2007-01-1461*, 2007.

[125] Karin Krüger, Lucas Ginzinger, and Heinz Ulbrich. Influences of leakage gap variations on the dynamics of hydraulic chain tensioners - experiment and simulation. *SAE Technical Paper 2008-01-0294*, 2008.

[126] Stefan Kurz. *Die numerische Behandlung elektromechanischer Systeme mit Hilfe der Kopplung der Methode der finiten Elemente und der Randelementmethode*. Fortschritt-Berichte VDI. VDI Verlag, 1998.

[127] Matthias Lang. *Berechnung und Optimierung von passiven permanentmagnetischen Lagern für rotierende Maschinen*. PhD thesis, Technische Universität Berlin, 2003.

[128] Günter Lautz. *Elektromagnetische Felder*. B. G. Teubner, 1969.

[129] E. W. Lee. Magnetostriction and magnetomechanical effects. *Reports on Progress in Physics*, 18:184–229, 1955.

[130] J. Lee and D.J. Thompson. Dynamic stiffness formulation, free vibration and wave motion of helical springs. *Journal of Sound and Vibration*, 239(2):297–320, 2001.

[131] Günther Lehner. *Elektromagnetische Feldtheorie*. Springer Verlag, 2003.

[132] R. I. Leine and Ch. Glocker. A set-valued force law for spatial Coulomb-Contensou friction. *European Journal of Mechanics*, 22:193–216, 2003.

[133] Klaus-Dieter Linsmeier and Achim Greis. *Elektromagnetische Aktoren*, volume 197 of *Die Bibliothek der Technik*. Verlag Moderne Industrie, 2007.

[134] Philipp Löbbert. *Möglichkeiten und Grenzen der Teillaststeuerung von Ottomotoren mit vollvariablem Ventilhub*. PhD thesis, Technische Universität Dresden, 2006.

[135] G. Maas. *Analyse des dynamischen Betriebsverhaltens von Ventiltrieben mit Hydrostößeln*. PhD thesis, RWTH Aachen, 1987.

[136] Eiji Maeno, Hiroshi Bunko, and Katsuhisa Yamaguchi. Development of an end-pivot type mechanical lash adjuster. *NTN Technical Review*, 75:78–85, 2007.

[137] Norberto Majlis. *The quantum theory of magnetism*. World Scientific, 2007.

[138] Hans-Jürgen Matthies. *Einführung in die Ölhydraulik*. B. G. Teubner, 1991.

[139] James Clerk Maxwell. *A treatise on electricity and magnetism*. Clarendon Press Oxford, 1873.

[140] Manfred Meissner and Hans-Jürgen Schorcht. *Metallfedern - Grundlagen, Werkstoffe, Berechnung und Gestaltung*. Springer Verlag, 1997.

[141] Günther Mendl. *Motordynamik und ihre Interaktion zu Festigkeit, Ladungswechsel und Akustik*. Number 500 in VDI Fortschritt-Berichte. VDI Verlag, 2002.

[142] Lothar Michalowsky and Jürgen Schneider. *Magnettechnik*. Vulkan Verlag, 2006.

[143] Hermann Middendorf, Jörg Theobald, Leonhard Lang, and Kai Hartel. Der 1,4-L-TSI-Ottomotor mit Zylinderabschaltung. *Motortechnische Zeitschrift*, 73(3):186–193, 2012.

[144] Michael Möller and Christop Glocker. Non-smooth modelling of electrical systems using the flux approach. *Nonlinear Dynamics*, 50:273–295, 2007.

[145] Jean Jacques Moreau. Unilateral contact and dry friction in finite freedom dynamics. In J. J. Moreau and P.D. Panagiotopoulos, editors, *International Center for Mechanical Science, Courses and Lectures*, 302, pages 1–82. Springer, 1988.

[146] John E. Mottershead. Finite elements for dynamical analysis of helical rods. *International Journal of Mechanical Sciences*, 22:267–283, 1980.

[147] John E. Mottershead. The large displacements and dynamic stability of springs using helical finite elements. *International Journal of Mechanical Science*, 24(9):547–558, 1982.

[148] MSC.Software Corporation. *ADAMS/Engine: Hydraulic Lash Adjuster Theory*, 2003. User Manual.

[149] Thomas H. Muhr. New technologies for engine valve springs. *SAE Technical Paper 930912*, 1993.

[150] Atsushi Murakami, Yoshitoshi Hagiwara, Takanori Igarashi, Toshio Ozone, Kazuo Matsuhashi, and Hiroaki Hayashi. Enhancement of automotive engines with β-Ti valve springs. *SAE Technical Paper 890470*, 890470, 1989.

[151] Kosuke Nagaya, Sadahiko Takeda, and Yoshitaka Nakata. Free vibration of coil springs of arbitrary shape. *International Journal for Numerical Methods in Engineering*, 23:1081–1099, 1986.

[152] Andreas Nicola. *Versuchsgestützte Dynamiksimulation hydraulisch gespannter Kettentriebe unter Drehungleichförmigkeiten*. PhD thesis, TU Kaiserslautern, 2008.

[153] Gustav Niemann. *Maschinenelemente - Band 1 Konstruktion und Berechnung von Verbindungen, Lagern, Wellen*. Springer Verlag, 1975.

[154] P. Niepage and P. Grindel. Ventilfedern mit Drahtquerschnitten, die von der Kreisform abweichen. *Draht*, 42(7/8):495–503, 1991.

[155] C. Ortmann and H. Skovbjerg. Powertrain analysis applications using ADAMS/Engine powered by FEV Part I: Valve spring. In *International ADAMS users conference*, Rome, 2000.

[156] Alexander Ostermann. A class of half-explicit runge-kutta methods for differential-algebraic systems of index 3. *Applied Numerical Mathematics*, 13:165–179, 1993.

[157] Giuseppe Peano. Démonstration de l'intégrabilité des équations différentielles ordinaires. *Mathematische Annalen*, 37(2):182–228, 1890.

[158] Linda Petzold. Differential/algebraic equations are not ode's. *SIAM Journal on Scientific and Statistical Computing*, 3(3):367–384, 1982.

[159] Linda Petzold and Per Lötstedt. Numerical solution of nonlinear differential equations with algebraic constraints II: Practical implications. *SIAM Journal on Scientific and Statistical Computing*, 7(3):720–733, 1986.

[160] Friedrich Pfeiffer. *Mechanical System Dynamics*, volume 40 of *Lecture Notes in Applied and Computational Mechanics*. Springer Verlag, 2008.

[161] Friedrich Pfeiffer and Christop Glocker. *Multibody Dynamics with Unilateral Contacts*. John Wiley Inc., 1996.

[162] J. W. Phillips and G. A. Costello. Large deflections of impacted helical springs. *The Journal of the Acoustic Society of America*, pages 967–973, 1971.

[163] P. J. Phlips, A. R. Schamel, and J. Meyer. An efficient model for valvetrain and spring dynamics. *SAE Technical Paper 890619*, 1989.

[164] Patrick Phlips and Andreas Schamel. The dynamics of valvetrains with hydraulic lash adjusters and the interaction with the gas exchange process. *SAE Technical Paper 910071*, 1991.

[165] A.P. Pisano and Ferdinand Freudenstein. An experimental and analytical investigation of the dynamic response of a high-speed cam follower system. Part 2: A combined, lumped/distributed parameter dynamic model. *Journal of Mechanisms Transmissions and Automation in Design*, 105:699–704, 1983.

[166] Anton Pischinger. *Die Steuerung der Verbrennungskraftmaschine*. Springer Verlag, 1948.

[167] P. Porot and J. Trapy. A numerical and experimental study of the effect of aeration of oil on valve trains equipped with hydraulic lash adjusters. *SAE Technical Paper 930997*, 1993.

[168] Joachim Post, Fredrik Borchsenius, and Friedrich Pfeiffer. *Simulation von Kettentrieben*. VDI Berichte Nr. 1467. VDI Verlag, 1999.

[169] L.F. Richardson. The deferred approach to the limit. *Phil. Trans., A*, 226:229–349, 1927.

[170] R. T. Rockafellar. Augmented Lagrangians and applications of the proximal point algorithm in convex programming. *Mathematics of Operations Research*, 1:97–116, 1976.

[171] Thomas Roschke, Sebastian Fraulob, René Seiler, and Thomas Bödrich. Bipolar magnetic actuators and approaches for their design. In *Proceedings of the 10th European Space Mechanisms and Tribology Symposium, 24-26 September 2003, San Sebastián, Spain*, pages 209–215. ESA Publications Division, 2003.

[172] Jörg Roß. *Das Betriebsverhalten des Ventiltriebs in 4-Takt-Verbrennungsmotoren unter Berücksichtigung des Einflusses von Nockenwellenschwingungen*. PhD thesis, RWTH Aachen, 1993.

[173] Jian Ruan and Richard Burton. Bulk modulus of air content oil in a hydraulic cylinder. In *ASME 2006 International Mechanical Engineering Congress and Exposition (IMECE2006)*, pages 259–269, 2006. Paper no. IMECE2006-15854.

[174] Ulf Säger. *Simulation des nichtlinearen dynamischen Verhaltens von Spiraldruckfedern bei stoßartiger Belastung*. PhD thesis, Universität Stuttgart, 1998.

[175] Otmar Franz Scharrer. *Einflusspotenzial variabler Ventiltriebe auf die Teillast-Betriebswerte von Saug-Ottomotoren - eine Studie mit Motorprozess-Simulation*. PhD thesis, Technischen Universität Berlin, 2005.

[176] Thorsten Schindler. *Spatial Dynamics of Pushbelt CVTs*. Number 730 in Fortschritt-Berichte VDI. VDI Verlag, 2010.

[177] Thorsten Schindler, Martin Förg, Markus Friedrich, Markus Schneider, Bastian Esefeld, Robert Huber, Roland Zander, and Heinz Ulbrich. Analysing dynamical phenomenons: Introduction to MBSim. In *1st Joint International Conference on Multibody System Dynamics, Lappeenranta, Finland*, 2010.

[178] Markus Schneider. *Hydaulische Systeme zur Phasenregelung von Ventiltrieben.* PhD thesis, TU München, 2012. to appear.

[179] Markus Schneider, Karin Krueger, and Heinz Ulbrich. Experiments and simulation of hydraulic cam phasing systems. *SAE Technical Paper 2008-01-1357*, 2008.

[180] Gerhard Schnell. *Magnete - Grundlagen, Aufbau und Anwendungen.* Verlag Karl Thiemig, 1973.

[181] Mark Sellnau and Eric Rask. Two-step variable valve actuation for fuel economy, emissions, and performance. *SAE Technical Paper 2003-01-0029*, 2003.

[182] Wolfgang Semet. *Entwicklung einer pneumatischen Ventilfeder für hochdrehende Serienmotoren.* PhD thesis, Universität Stuttgart, 2007.

[183] Takasuke Shikida, Yoshikatsu Nakamura, Tamio Nakakubo, and Hiroyuki Kawase. Development of the high speed 2ZZ-GE engine. *SAE Technical Paper 2000-01-0671*, 2000.

[184] Friedrich-Wilhelm Speckens. *Optimierungsstrategien für die Auslegung von Ventilfedern in Tassenstößelventiltrieben.* PhD thesis, Technische Hochschule Aachen, 1994.

[185] A. Stiegelmeyr. *Zur numerischen Berechnung strukturvarianter Systeme.* PhD thesis, TU München, Düsseldorf, 2001.

[186] A. Stiegelmeyr and F. Pfeiffer. A Time Stepping Algorithm for Mechanical Systems with Unilateral Contacts. In *Proc. of the ASME Design Engineering Technical Conferences*, Las Vegas, Nevada, USA, 1999.

[187] J. Stoer and R. Bulirsch. *Introduction to Numerical Analysis. Third Edition.* Texts in Applied Mathmatics. Springer Verlag, 2002.

[188] C. Studer, R. I. Leine, and Ch. Glocker. Step size adjustment and extrapolation for time-stepping schemes in non smooth mechanics. *Int. J. Numer. Meth. Engng*, 2008.

[189] Radek Tichanek, David Fremut, and Robert Cihak. The over-head cam (ohc) valve train computer model. SIMPACK User Meeting, 2006.

[190] O. Tietjens. *Strömungslehre.* Zweiter Band, Bewegung der Flüssigkeiten und Gase. Springer Verlag, 1970.

[191] George E. Totten, editor. *Handbook of Hydraulic Fluid Technology.* CRC Press, 1999.

[192] Svent Tröster. Functional test of an variable valve train system. unpublished diploma thesis, Audi AG, 2005.

[193] Erich A. Truckenbrodt. *Fluidmechanik*, volume 1. Springer Verlag, 1980.

[194] C.A. Truesdell and W. Noll. *The non-linear field theories of mechanics.*, volume 3 of *Encyclop. of Physics*. Springer Verlag, 1965.

[195] L. Ubbelohde. *Zur Viskosimetrie*. Hirzel Verlag, 7 edition, 1965.

[196] Heinz Ulbrich. Comparison of different actuator concepts for applications in rotating machinery. *International Journal of Rotating Machinery*, 1(1):61–71, 1994.

[197] Heinz Ulbrich. *Maschinendynamik*. B. G. Teubner, Stuttgart, 1996.

[198] Harald Unger, Christian Schwarz, Jürgen Schneider, and Karl-Friedrich Koch. Valvetronic - Erfahrungen aus sieben Jahren Großserie und Ausblick in die Zukunft. *Motortechnische Zeitschrift*, 69(7/8):598–605, 2008.

[199] Richard van Basshuysen and Fred Schäfer, editors. *Handbuch Verbrennungsmotor: Grundlagen, Komponenten, Systeme, Perspektiven*. Springer Vieweg, 2011.

[200] Wilhelm R. A. Vauck and Hermann A. Müller. *Grundoperationen chemischer Verfahrenstechnik*. Wiley VCH Verlag, 2000.

[201] C. Walther. Anforderungen an Schmiermittel. *Zeitschrift für den Maschinenbau*, 19:671, 1931.

[202] Tobias Wittkopp. *Mehrkörpersimulation von Schraubendruckfedern*. Verlag ISLE, Illmenau, 2005.

[203] William Henry Wittrick. On elastic wave propagation in helical springs. *International Journal of Mechanical Sciences*, 8:25–47, 1966.

[204] Peter Wriggers. *Nichtlineare Finite-Element-Methode*. Springer Verlag, 2001.

[205] Rainer Wurms, Stefan Dengler, Ralf Budack, Günther Mendl, Thomas Dicke, and Alex Eiser. Audi valvelift system - ein neues innovatives Ventiltriebssystem von Audi. In *15. Aachener Kolloquium - Fahrzeug- und Motorentechnik*, 2006.

[206] Roland Zander, Thorsten Schindler, Markus Friedrich, Robert Huber, Martin Förg, and Heinz Ulbrich. Non-smooth dynamics in academia and industry: recent work at TU München. *Acta Mechanica*, 195(1-4):167–183, 2008.

[207] Dequan Zou and Harold E. McCormick. Dynamic model and computer simulation of valve train assemblies with hydraulic lash adjuster. *SAE Technical Paper 960351*, 1996.

[208] Florian Zschunke. *Aktoren auf Basis des magnetorheologischen Effekts*. PhD thesis, Universität Erlangen-Nürnberg, 2005.