

Werner Schiehlen  
*Editor*



International Centre  
for Mechanical Sciences

# Dynamical Analysis of Vehicule Systems

CISM Courses and Lectures, vol. 497



SpringerWienNewYork

 SpringerWienNewYork

# CISM COURSES AND LECTURES

Series Editors:

The Rectors

Giulio Maier - Milan

Jean Salençon - Palaiseau

Wilhelm Schneider - Wien

The Secretary General

Bernhard Schrefler - Padua

Executive Editor

Paolo Serafini - Udine

The series presents lecture notes, monographs, edited works and proceedings in the field of Mechanics, Engineering, Computer Science and Applied Mathematics.

Purpose of the series is to make known in the international scientific and technical community results obtained in some of the activities organized by CISM, the International Centre for Mechanical Sciences.

INTERNATIONAL CENTRE FOR MECHANICAL SCIENCES

COURSES AND LECTURES - No. 497



**DYNAMICAL ANALYSIS  
OF VEHICLE SYSTEMS**  
**THEORETICAL FOUNDATIONS  
AND ADVANCED APPLICATIONS**

EDITED BY

WERNER SCHIEHLEN  
UNIVERSITY OF STUTTGART, GERMANY

**SpringerWienNewYork**

This volume contains 226 illustrations

This work is subject to copyright.  
All rights are reserved,  
whether the whole or part of the material is concerned  
specifically those of translation, reprinting, re-use of illustrations,  
broadcasting, reproduction by photocopying machine  
or similar means, and storage in data banks.

© 2007 by CISM, Udine

Printed in Italy

SPIN 12068067

All contributions have been typeset by the authors.

ISBN 978-3-211-76665-1 SpringerWienNewYork

## PREFACE

*This book contains an edited version of the lecture notes used for the Course "Dynamical Analysis of Vehicle Systems - Theoretical Foundations and Advanced Applications" offered at the Centre International des Sciences Mécaniques (CISM) during the Fiszdon Session. The Course took place in CISM's Palazzo del Torso in Udine, Italy, 23 - 27 October 2006. The Course was well attended by engineers from academia and industry, with a total number of 45 persons from twelve countries.*

*This volume presents an integrated approach to the common fundamentals of rail and road vehicles based on multibody system dynamics, rolling wheel contact and control system design. The mathematical methods presented allow an efficient and reliable analysis of the resulting state equations, and may also be used to review simulation results from commercial vehicle dynamics software.*

*The book will also provide a better understanding of the basic physical phenomena of vehicle dynamics most important for the engineering practice in research and in industry. Particular attention will be paid to developments of future road and rail vehicles. Again, mechatronic trains and mechatronic cars show many similarities which result in an interdisciplinary stimulation of the design concepts used. The automation of individual vehicle traffic on roads, and on rail, is an important point of issue in the future: Drivers reading the newspaper, watching television, surfing the Internet while their vehicles automatically find their way to the desired destination - rapid and secure. The course features two recent developments. The Railcab is an individual vehicle on existing railway tracks with point-to-point link in a complex controlled network. The Driver Assistance Systems are devices to control distance and keep in lane on existing roads, relieving the driver and leading to a significant improvement in driving safety and comfort.*

*The Course was originally initiated together with Professor Karl Popp, University of Hanover, Germany who passed away unexpectedly. The Lecturers and the Editor of this volume agreed to continue and, thus, the Course was delivered in memoriam of Professor Karl Popp, too. We thank Professor Giulio Maier, Rector of CISM, for the kind support during the final coordination of the Course. Further, the CISM staff is acknowledged for the excellent organisation. Finally, Professor Paolo Serafini is gratefully acknowledged for his encouragement to publish these lecture notes and his patience while it took longer to complete their editing in book form.*

*Werner Schiehlen*

## CONTENTS

Vehicle and Guideway Modelling: Suspension Systems <i>by W. Schiehlen</i> .....	1
Dynamics of Railway Vehicles and Rail/Wheel Contact <i>by H. True</i> .....	75
Tyre Models, Propulsion and Handling of Road Vehicles <i>by P. Lugner</i> .....	129
Dynamics of Motorcycles: Stability and Control <i>by R. Sharp</i> .....	183
Control of Rail Vehicles: Mechatronic Technologies <i>by R. Goodall</i> .....	231
RailCab System: Engineering Aspects <i>by J. Lückel et al.</i> .....	237
Advanced Road Vehicles: Control Technologies, Driver Assistance <i>by D. Ammon and W. Schiehlen</i> .....	283

# Vehicle and Guideway Modelling: Suspensions Systems

Werner Schiehlen

Institute of Engineering and Computational Mechanics, University of Stuttgart, Germany

**Abstract** Performance, safety and comfort of a vehicle are related to its low frequency motions. The corresponding mechanical models are characterized for all kinds of vehicles by stiff parts represented as rigid bodies and soft components like springs, dampers and actuators. The method of multibody systems is most appropriate for the analysis of vehicle motions and vibrations up to 50 Hz. In this contribution the derivation of the equations of motions of multibody systems is shown step by step up to the computer-aided evaluation of these equations.

Starting with kinematics for rigid body vehicle systems, the foundations of dynamics together with the principles of d'Alembert and Jourdain are used to get the equations of motion. Then, some aspects of multibody dynamics formalisms and computer codes for vehicle dynamics are discussed. Further, models of randomly uneven guideways are presented. Performance criteria for ride comfort and safety are considered. Finally, the analysis of the suspension of a car model is presented in detail.

## 1 Kinematics

The elements of multibody systems for vehicle modelling, see Figure 1, include rigid bodies which may also degenerate to particles, coupling elements like springs, dampers or force controlled actuators as well as ideal, i.e. rigid kinematical connecting elements like joints, bearings, rails and motion controlled actuators. The coupling and connection elements are generating internal forces and torques between the bodies of the system and external forces with respect to the environment. Both of them are considered as massless elements. The kinematical constraints resulting from the connecting elements may be holonomic or nonholonomic, scleronomic or rheonomic, respectively. Holonomic constraints reduce the motion space of the system while nonholonomic constraints reduce the velocity space in addition. The constraint equations are called rheonomic if they depend explicitly on time, and scleronomic otherwise. Real vehicle systems are subject to holonomic constraints only which may be given by geometrical or integrable kinematical conditions. However, in more simplified models, e.g. rolling of a rigid wheel or wheelset on a rigid plane, nonholonomic constraints may occur. Some configurations of holonomic connecting elements are listed in Table 1 depending on the number of degrees of freedom characterizing the remaining possibilities of motion. Now the motion of vehicle parts will be described mathematically depending on space and time. This is the task of kinematics.



**Table 1.** Configurations of holonomic connecting elements

Motion	Degrees of Freedom		
	1	2	3
Rotary	Revolute Joint	Universal Joint	Spherical Joint
Linear	Prismatic Joint	Planar Joint	
Mixed	Screw Joint	Cylindric Joint	General Planar Joint

### 1.1 Frames of Reference for Vehicle Kinematics

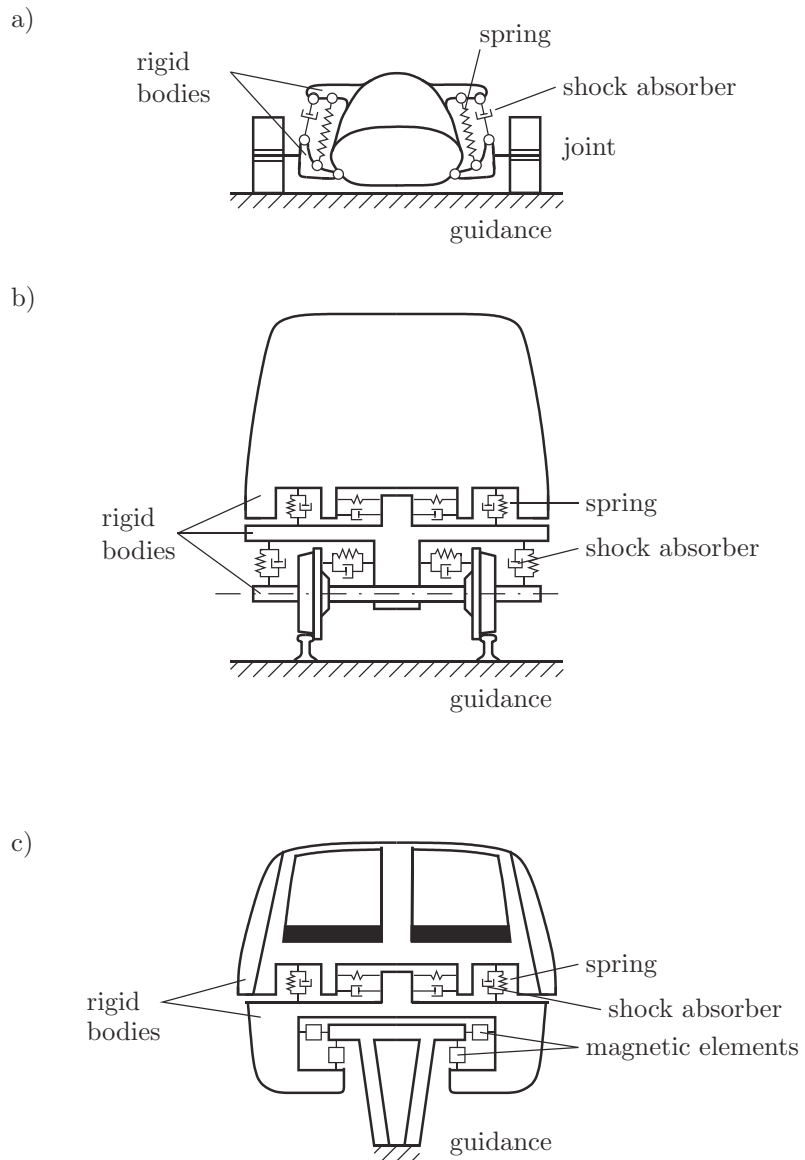
A prerequisite for the mathematical description of position, velocity and acceleration of a mechanical system is the definition of appropriate frames of reference. The frames required in vehicle dynamics are shown in Figure 2 with the details summarized in Table 2. There will be used only right-handed Cartesian frames with the unit base vectors  $\mathbf{e}_v$ ,  $|\mathbf{e}_v| = 1$  where the Greek indices generally take the integers 1, 2, 3. A basis or frame  $\{O, \mathbf{e}_v\}$ , respectively, is completely defined by its origin  $O$  and its base vectors  $\mathbf{e}_v$ . For distinction between different frames the upper right index is used if necessary. The inertial frame  $\{O^I, \mathbf{e}_v^I\}$  serves as the general reference frame, in particular for the evaluation of the acceleration. The given trajectory of the vehicle is assumed to be a space curve with the moving frame  $\{O^B, \mathbf{e}_v^B\}$  also known as Frenet frame or moving trihedron. The origin  $O^B$  is moving with a given speed tangential to the trajectory. The reference frame  $\{O^R, \mathbf{e}_v^R\}$  is closely related to the moving frame. Its origin and the first unit vector coincide with the moving frame  $O^R = O^B$ ,  $\mathbf{e}_1^R = \mathbf{e}_1^B$ . The second base vector  $\mathbf{e}_2^R$ , however, is parallel to the guideway surface considering the bank of the road or the track, respectively, pointing to the right with respect to the direction of motion.

The body-fixed frame  $\{O^i, \mathbf{e}_v^i\}$  is the principal axis frame of the rigid body  $K_i$  located in its center of mass  $C_i$ . This frame describes uniquely the position in space of the body. Finally, there is defined a local frame  $\{O^j, \mathbf{e}_v^j\}$  to describe constraint elements between bodies. It is oriented according to the local specifications like the direction of a joint axis. In the following a frame is simply identified by its name (upper right index) only.

### 1.2 Kinematics of a Rigid Body in an Inertial Frame

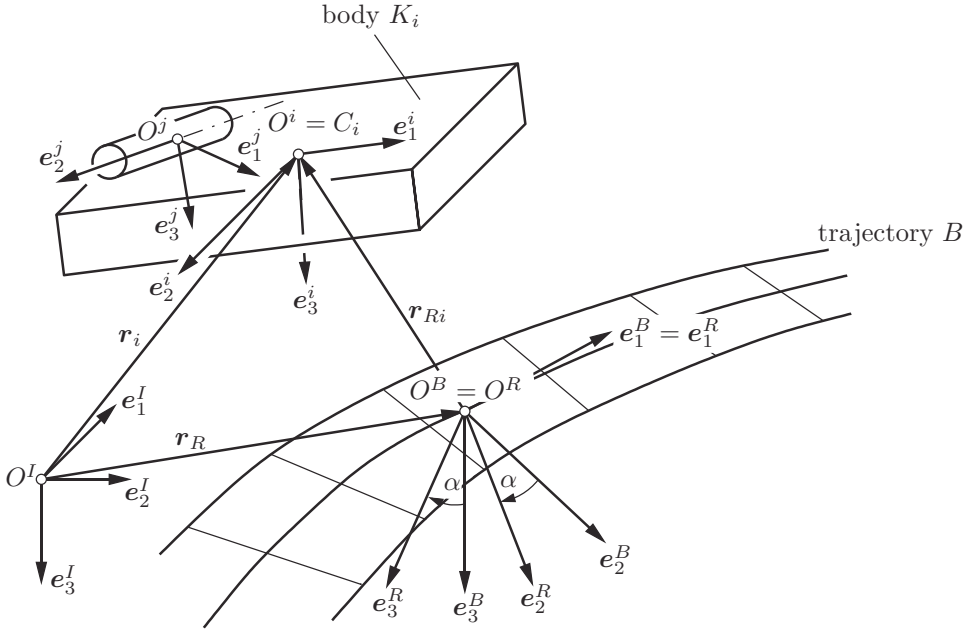
First of all some definitions and remarks on the nomenclature are presented. The position of a particle  $P$  in space is uniquely defined by the position vector  $\mathbf{x}$  represented in the inertial frame  $\{O^I, \mathbf{e}_v^I\}$  by its coordinates  $x_v$  as

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3 . \quad (1.1)$$



**Figure 1.** Multibody models of vehicles:

a) road vehicle, b) rail vehicle, c) magnetically levitated vehicle



**Figure 2.** Frames of reference

This set of coordinates may be summarized in a column matrix  $\mathbf{x}^I$  often simply called a vector, i.e.,

$$\mathbf{x}^I = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \equiv [x_1 \quad x_2 \quad x_3]^T \quad (1.2)$$

where upper right index defines the frame in which the coordinates are measured. This index will often be deleted if there isn't any possibility for a mix-up of frames or if there is used only one frame identified in the text. The goal is to present all vector and tensor quantities in one common frame, e.g. the inertial frame  $I$ . Then, it is possible, to integrate subsystems easily into the complete system.

For a particle  $P$  moving in time its coordinates are time-dependent, too, and they define a trajectory in space. The mathematical representation results in the vector equation  $\mathbf{x} = \mathbf{x}(t)$  equivalent to three scalar equations according to the three degrees of freedom of the particle in the three-dimensional space. The velocity  $\mathbf{v}(t)$  and the acceleration  $\mathbf{a}(t)$  of the particle follow by differentiation with respect to time as

$$\mathbf{v}(t) = \frac{d^I \mathbf{x}(t)}{dt}, \quad \mathbf{v}^I(t) = \dot{\mathbf{x}}^I(t) = [\dot{x}_1 \quad \dot{x}_2 \quad \dot{x}_3]^T, \quad (1.3)$$

$$\mathbf{a}(t) = \frac{d^I \mathbf{v}(t)}{dt}, \quad \mathbf{a}^I(t) = \dot{\mathbf{v}}^I(t) = \ddot{\mathbf{x}}(t) = [\ddot{x}_1 \quad \ddot{x}_2 \quad \ddot{x}_3]^T. \quad (1.4)$$

**Table 2.** Frames of reference

Frame of reference	Origin of frame	Orientation of axis
Inertial frame $\{O^I, \mathbf{e}_v^I\}$	$O^I$ space fixed	$\mathbf{e}_1^I, \mathbf{e}_2^I$ in horizontal plane $\mathbf{e}_3^I$
Moving trihedron $\{O^B, \mathbf{e}_v^B\}$	$O^B$ trajectory-fixed	$\mathbf{e}_1^B \equiv \mathbf{e}_t$ tangential to trajectory $\mathbf{e}_2^B \equiv \mathbf{e}_n$ normal to trajectory $\mathbf{e}_3^B \equiv \mathbf{e}_b$ bi-normal to trajectory
Reference frame $\{O^R, \mathbf{e}_v^R\}$	$O^R$ trajectory-fixed	$\mathbf{e}_1^R \equiv \mathbf{e}_1$ $\mathbf{e}_2^R$ in guideway plane $\mathbf{e}_3^R$ normal to guideway plane
Body-fixed frame $\{O^i, \mathbf{e}_v^i\}$	$O^i \equiv C_i$ body-fixed in center of mass	$\mathbf{e}_v^i$ principle inertia axes
Local frame $\{O^j, \mathbf{e}_v^j\}$	$O^j$ arbitrary	$\mathbf{e}_v^j$ locally specified axes

The upper right index refers to the frame of reference in which the operations, in particular the differentiation, have to be executed. In the inertial frame  $I$  the differentiation of vectors is just performed by differentiation of the scalar coordinates.

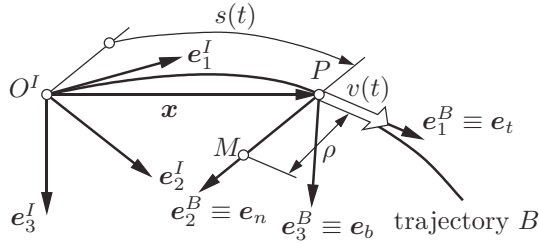
The motion of a particle  $P$  on a curvilinear trajectory in space may be shown in the moving frame  $B$ , too. The position of the point is uniquely identified by the arc length  $s(t)$  as a generalized coordinate, Figure 3. Then, the position vector  $\mathbf{r}$  is a function of the arc length,  $\mathbf{r} = \mathbf{r}(s)$ . For the velocity and acceleration vector it yields, see e.g. Magnus and Müller (1990),

$$\begin{aligned} \mathbf{v}(t) &= \frac{d^I \mathbf{r}(t)}{dt} = \frac{d^I \mathbf{r}(s)}{ds} \frac{ds}{dt} = v \mathbf{e}_t, \quad v = \dot{s}, \\ \mathbf{v}^B(t) &= \begin{bmatrix} \dot{s} & 0 & 0 \end{bmatrix}^T, \end{aligned} \quad (1.5)$$

$$\begin{aligned} \mathbf{a}(t) &= \frac{d^I \mathbf{v}(t)}{dt} = a_t \mathbf{e}_t + a_n \mathbf{e}_n = \dot{v} \mathbf{e}_t + \frac{v^2}{\rho} \mathbf{e}_n, \\ \mathbf{a}^B(t) &= \begin{bmatrix} \ddot{s} & s^2/\rho & 0 \end{bmatrix}^T, \end{aligned} \quad (1.6)$$

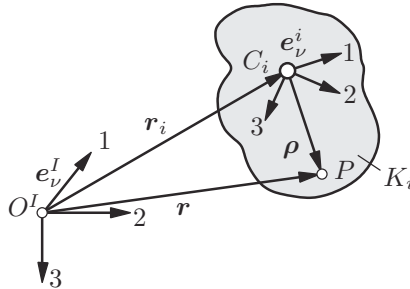
where  $\rho$  is the curvature of the trajectory in point  $P$ . Further, it is

$v = \dot{s}$  the tangent velocity,  
 $a_t = \dot{v} = \ddot{s}$  the tangent acceleration, and  
 $a_n = v^2/\rho = \dot{s}^2/\rho$  the normal or centripetal acceleration.



**Figure 3.** Trajectory of particle  $P$

Special cases of the general motion in space are the motion in a straight line ( $\rho \rightarrow \infty$ ), and the motion on a circle ( $\rho = \text{const}$ ). Often the functions  $s(t)$ ,  $\dot{s}(t)$ ,  $\ddot{s}(t)$ ,  $\dot{s}(s)$ ,  $\ddot{s}(s)$  are depicted in kinematical diagrams for graphical visualization of the motion along a track.



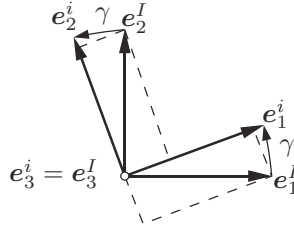
**Figure 4.** Position of a rigid body  $K_i$  in the inertial frame  $I$

The translational motion of a rigid body  $K_i$  is completely described by the general relations for a particle applied to a body-fixed point, e.g. the center of mass  $C_i$  of the rigid body, and the corresponding position vector  $\mathbf{r}_i$ , see Figure 4. The rotational motion of a rigid body  $K_i$  follows from the relative position of two frames where one of them is a body-fixed frame. For coinciding origins the position of the body-fixed frame  $i$  relative to the inertial frame  $I$  is uniquely defined by three rotation angles according to the three rotational degrees of freedom of a rigid body in space. Both frames are related to each other by three elementary rotations performed successively around different base vectors using three rotation angles. If, for example, the frame  $i$  is revolved around the coinciding 3-axes of frame  $I$  and  $i$  by the angle  $\gamma$ , the relation of the corresponding base vectors is

given by the matrix  $\mathbf{S}^{Ii}$  as shown in Figure 5 and reads as

$$\begin{bmatrix} \mathbf{e}_1^I \\ \mathbf{e}_2^I \\ \mathbf{e}_3^I \end{bmatrix} = \underbrace{\begin{bmatrix} \cos \gamma & -\sin \gamma & 0 \\ \sin \gamma & \cos \gamma & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\mathbf{S}^{Ii} = \gamma_3} \begin{bmatrix} \mathbf{e}_1^i \\ \mathbf{e}_2^i \\ \mathbf{e}_3^i \end{bmatrix} \quad (1.7)$$

The row  $\nu$  of the elementary rotation matrix  $\gamma_3$  is composed of the coordinates of the



**Figure 5.** Elementary rotation with angle  $\gamma$  around 3-axis

base vector  $\mathbf{e}_\nu^I$  in frame  $i$ . The corresponding matrices for positive rotations around the remaining axes read as

$$\boldsymbol{\alpha}_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha \\ 0 & \sin \alpha & \cos \alpha \end{bmatrix}, \quad \boldsymbol{\beta}_2 = \begin{bmatrix} \cos \beta & 0 & \sin \beta \\ 0 & 1 & 0 \\ -\sin \beta & 0 & \cos \beta \end{bmatrix} \quad (1.8)$$

where the elementary rotation matrices are characterized by the name of the rotation angle while the index defines the axis of rotation. There are numerous possibilities to choose the name of the angle and the sequence of the rotation axes used which is not commutative. In vehicle dynamics the Cardan angles  $\alpha, \beta, \gamma$  are often used, see Figure 6, which are different from the well-known Euler angles  $\psi, \vartheta, \varphi$ .

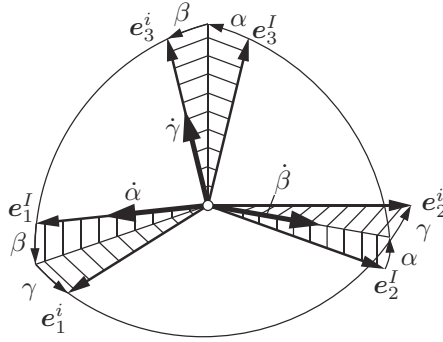
The resulting rotation matrices  $\mathbf{S}^{Ii}$  which present the relation between frames  $I$  and  $i$  are obtained by the corresponding matrix multiplications

$$\mathbf{S}^{Ii}(\alpha, \beta, \gamma) = \boldsymbol{\alpha}_1 \boldsymbol{\beta}_2 \gamma_3, \quad \mathbf{S}^{Ii}(\psi, \vartheta, \varphi) = \boldsymbol{\psi}_3 \boldsymbol{\vartheta}_1 \boldsymbol{\varphi}_3. \quad (1.9)$$

Since matrix products are not commutative, the sequence of the elementary rotations has to be strongly observed. The Cardan and Euler angles, respectively, are defined by successive rotations around the 1-, 2-, 3-axis and 3-, 1-, 3-axis, respectively, starting from the inertial frame  $I$ . The sequence of the elementary rotations is uniquely identified by the sequence of the indices of the elementary rotation matrices as shown in (1.9).

The rotation matrices are orthogonal matrices

$$\mathbf{S}^{iI} \left( \mathbf{S}^{Ii} \right)^{-1} = \left( \mathbf{S}^{Ii} \right)^T = \mathbf{S}^{iI}, \quad \det \mathbf{S} = +1 \quad (1.10)$$



**Figure 6.** Spatial rotation with Cardan angles  $\alpha$ ,  $\beta$ ,  $\gamma$

where the inversion is also represented by the exchange of the upper indices. The inverse rotation matrix is simply found by transposition of the original rotation matrix  $\mathbf{S}^{Ii}$ . Using Cardan angles the rotation matrix reads explicitly as

$$\mathbf{S}^{Ii}(\alpha, \beta, \gamma) = \begin{bmatrix} c\beta c\gamma & -c\beta s\gamma & s\beta \\ c\alpha s\gamma + s\alpha s\beta c\gamma & c\alpha c\gamma - s\alpha s\beta s\gamma & -s\alpha c\beta \\ s\alpha s\gamma - c\alpha s\beta c\gamma & s\alpha c\gamma + c\alpha s\beta s\gamma & c\alpha c\beta \end{bmatrix} \quad (1.11)$$

where the abbreviations  $c$  and  $s$  stands for  $\cos$  and  $\sin$ , respectively. In applications often small rotations are found,  $\alpha, \beta, \gamma \ll 1$ , resulting in the linearized rotation matrix

$$\mathbf{S}^{Ii}(\alpha, \beta, \gamma) = \begin{bmatrix} 1 & -\gamma & \beta \\ \gamma & 1 & -\alpha \\ -\beta & \alpha & 1 \end{bmatrix}. \quad (1.12)$$

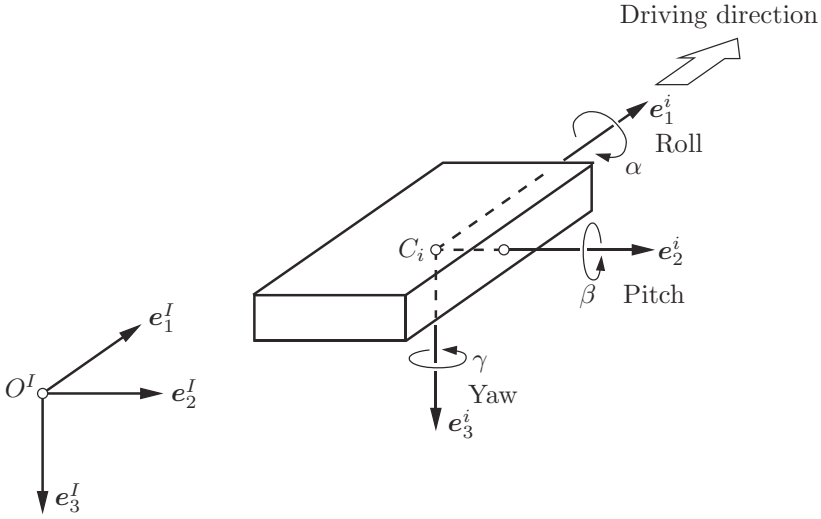
This result is also obtained if the elementary rotation matrices are linearized and multiplied with each other. Due to the vector property of small rotations the sequence of the multiplications has no longer to be considered. For small rotations the Cardan angles may be assigned directly to the rotational motions around the body-fixed axes, Figure 7. In vehicle engineering the following notations are used,

- $\alpha$  roll motion,
- $\beta$  pitch motion,
- $\gamma$  yaw motion.

The coordinates of a vector  $\mathbf{x}$  read differently for different frames. The relation between the coordinates  $\mathbf{x}^i$  in frame  $i$  and the coordinates  $\mathbf{x}^I$  in frame  $I$  is given by the transformation law for vector coordinates as

$$\mathbf{x}^i = \mathbf{S}^{iI} \mathbf{x}^I \quad \text{and} \quad \mathbf{x}^I = \mathbf{S}^{Ii} \mathbf{x}^i, \quad (1.13)$$

respectively, what is easily proven by (1.10). Please observe that the same indices appear in both forms in a neighbouring sequence. This property is often helpful in applications.



**Figure 7.** Notation of rotary vehicle motions

Further, it has to be pointed out that the rotation matrix  $\mathbf{S}$  is a function of time,  $\mathbf{S} = \mathbf{S}(t)$ , what has to be considered for time derivations.

The position of a rigid body  $K_i$  in the inertial frame is uniquely described by the position quantities  $\{\mathbf{r}_i, \mathbf{S}^{Ii}\}$  which characterize the body-fixed frame  $\{C^i, \mathbf{e}^i\}$ . During motion the position quantities are functions of time. Thus, the position coordinates of an arbitrary particle  $P$  of the rigid body read in the inertial frame  $I$  as

$$\mathbf{r}^I(t) = \mathbf{r}_i^I(t) + \boldsymbol{\rho}^I(t), \quad \boldsymbol{\rho}^I(t) = \mathbf{S}^{Ii}(t)\boldsymbol{\rho}^i, \quad (1.14)$$

where in the body-fixed frame it yields  $\boldsymbol{\rho}^i = \mathbf{const}$ , see also Figure 4.

The motion of a rigid body  $K_i$  will be now presented in the inertial frame  $I$ , too. The change of the position of its particle  $P$  with respect to time relative to frame  $I$  is found by differentiation of (1.14) as

$$\dot{\mathbf{r}}^I(t) = \dot{\mathbf{r}}_i^I(t) + \dot{\mathbf{S}}^{Ii}(t)\boldsymbol{\rho}^i = \dot{\mathbf{r}}_i^I(t) + \dot{\mathbf{S}}^{Ii}(t)\mathbf{S}^{iI}(t)\boldsymbol{\rho}^I(t). \quad (1.15)$$

The first term on the right-hand side represents the translational velocity of the origin  $C_i$  of the body-fixed frame  $i$ . The second term is obviously related to the rotation of the body-fixed frame and represents the body's rotation. This term will now be discussed in more detail. The matrix product  $[\dot{\mathbf{S}}(t)\mathbf{S}^T(t)]$  is screw symmetric, i.e.,  $[\bullet] = -[\bullet]^T$ , what follows immediately from the differentiation of the orthogonality condition  $\mathbf{S}(t)\mathbf{S}^T(t) = \mathbf{E}$  according to (1.10):

$$\begin{aligned} \frac{d}{dt} [\mathbf{S}(t)\mathbf{S}^T(t)] &= \dot{\mathbf{S}}(t)\mathbf{S}^T(t) + \mathbf{S}(t)\dot{\mathbf{S}}^T(t) \\ &= \dot{\mathbf{S}}(t)\mathbf{S}^T(t) + [\dot{\mathbf{S}}(t)\mathbf{S}^T(t)]^T = \mathbf{0}. \end{aligned} \quad (1.16)$$



The matrix product  $[\bullet]$  will be abbreviated by the symbol  $\tilde{\omega}(t)$  and identified by the corresponding three coordinates  $\omega_v = \omega_v(t)$  as follows

$$\dot{\mathbf{S}}^{Ii} (\mathbf{S}^{Ii})^T = \dot{\mathbf{S}}^{Ii} \mathbf{S}^{iI} = \tilde{\omega}_{Ii}^I(t) := \begin{bmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{bmatrix}, \quad \omega_{Ii}^I = \begin{bmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{bmatrix}. \quad (1.17)$$

Both quantities, the screw symmetric tensor  $\tilde{\omega}_{Ii}^I$  and the corresponding rotational velocity vector  $\omega_{Ii}^I$ , respectively, describe the rotational motion of system  $i$  or body  $K_i$ , respectively, relative to the inertial frame  $I$ . The upper indices indicate that both quantities are represented in the inertial frame  $I$ . If there is no chance for mixing up the frames the upper and lower index  $I$  is simply deleted. The screw symmetric tensor corresponding with a vector  $(\bullet)$  is identified by the symbol  $(\tilde{\bullet})$  and it replaces the vector product

$$\tilde{\omega}\rho \equiv \omega \times \rho. \quad (1.18)$$

In coordinates, in any frame, one gets accordingly

$$\tilde{\omega}\rho = \begin{bmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{bmatrix} \begin{bmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \end{bmatrix} = \begin{bmatrix} \omega_2\rho_3 - \omega_3\rho_2 \\ \omega_3\rho_1 - \omega_1\rho_3 \\ \omega_1\rho_2 - \omega_2\rho_1 \end{bmatrix}. \quad (1.19)$$

This notation of the vector product is most valuable for numerical computations since the vector product is not defined in matrix calculus.

The rotational velocity vector  $\omega_{Ii}^i$  in the body-fixed frame  $i$  follows from transformation or direct evaluation, respectively. The application of transformation (1.13) to (1.19) results in

$$\tilde{\omega}_{Ii}^i = \mathbf{S}^{iI} \tilde{\omega}_{Ii}^I \mathbf{S}^{Ii} = \mathbf{S}^{iI} \left( \dot{\mathbf{S}}^{Ii} \mathbf{S}^{iI} \right) \mathbf{S}^{Ii} = \mathbf{S}^{iI} \dot{\mathbf{S}}^{Ii} = \left( \mathbf{S}^{Ii} \right)^T \dot{\mathbf{S}}^{Ii}. \quad (1.20)$$

The first and second term of (1.20) represents the transformation law for tensor coordinates where the same indices appear again in a neighbouring sequence, the first and last term show the direct evaluation. The vector corresponding to (1.20) is  $\omega_{Ii}^i = \mathbf{S}^{iI} \omega_{Ii}^I$  where the transformation law for vector coordinates has been used again.

From (1.15) and (1.17) it follows for rigid body  $K_i$

$$\mathbf{v}^I(t) = \mathbf{v}_i^I(t) + \tilde{\omega}_{Ii}^I(t) \rho^I(t). \quad (1.21)$$

Considering (1.18), this is the relation for rigid body kinematics well-known from each mechanics textbook as

$$\mathbf{v}(t) = \mathbf{v}_i(t) + \omega_{Ii}(t) \times \rho(t). \quad (1.22)$$

The relations (1.21) or (1.22), respectively, represent the motion of a rigid body composed by an absolute translational velocity  $\mathbf{v}_i$  of the body-fixed reference point  $O^i = C^i$ , and a rotation with the angular velocity  $\omega_{Ii}$ . The fundamental kinematical quantities  $\{\mathbf{v}_i, \omega_{Ii}\}$  are also denoted as twist characterizing uniquely the motion of a rigid body.

### 1.3 Kinematics of a Rigid Body in a Moving Reference Frame

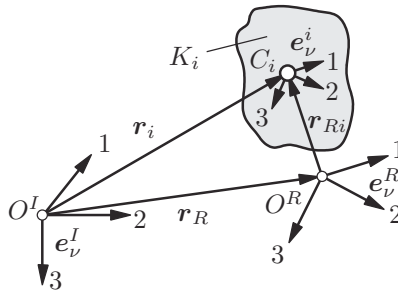
From a mathematical point of view the kinematical description of the motion of a rigid body is most convenient in the inertial frame  $I$  resulting in a more simple representation of the fundamental laws of mechanics. In engineering applications, however, a moving reference frame  $R$  related to the vehicle or the guideway, respectively, turns out to be more adequate. The frame  $R$  allows a problem-oriented choice of the coordinates and an efficient description of the forces and torques acting on the system. Moving reference frames are also useful in experiments since many measurement data are not related to the inertial frame. The choice of the reference frame  $R$  depends on the problem under consideration. In many cases the frame  $R$  characterizes the large nonlinear reference motion of a vehicle while the small deviation from the reference motion results in linear kinematical relations even for rotations.

In the following the motion of a rigid body is represented in a moving reference frame  $R$  the motion of which is known in the inertial frame  $I$  by the position vector  $\mathbf{r}_R(t)$  of its origin, and rotation matrix  $\mathbf{S}^{IR}(t)$ , see Figure 8. This means that the translational and angular guidance velocities are also known,  $\mathbf{v}_R^I = \dot{\mathbf{r}}_R^I$ ,  $\tilde{\boldsymbol{\omega}}_{IR}^I = \dot{\mathbf{S}}^{IR}\mathbf{S}^{RI}$  according to (1.3) and (1.17). Considering Figure 8, the absolute position quantities  $\{\mathbf{r}_i, \mathbf{S}^{Ii}\}$  of the rigid body reads as

$$\mathbf{r}_i^I(t) = \mathbf{r}_R^I(t) + \mathbf{S}^{IR}(t)\mathbf{r}_{Ri}^R(t), \quad (1.23)$$

$$\mathbf{S}^{Ii}(t) = \mathbf{S}^{IR}(t)\mathbf{S}^{Ri}(t). \quad (1.24)$$

The absolute motion  $\{\mathbf{v}_i, \boldsymbol{\omega}_{Ii}\}$  of the rigid body  $K_i$  is now found by formal differentiation



**Figure 8.** Position of a rigid body in the reference frame  $R$

in frame  $I$  and subsequent transformation in frame  $R$  as

$$\mathbf{v}_i^R(t) = \mathbf{v}_R^R(t) + \tilde{\boldsymbol{\omega}}_{IR}^R(t)\mathbf{r}_{Ri}^R(t) + \dot{\mathbf{r}}_{Ri}^R(t), \quad (1.25)$$

$$\boldsymbol{\omega}_{Ii}^R = \boldsymbol{\omega}_{IR}^R + \boldsymbol{\omega}_{Ri}^R. \quad (1.26)$$

Due to the rotation of the frame  $R$  the guidance motion and the relative motion characterized by the indices  $IR$  and  $Ri$ , respectively, are not simply added but there appears

an additional term in (1.25). This implies the well-known law of differentiation in a rotating frame

$$\frac{d^I}{dt} \mathbf{r}(t) = \frac{d^R}{dt} \mathbf{r}(t) + \boldsymbol{\omega}_{IR}(t) \times \mathbf{r}(t). \quad (1.27)$$

By formal differentiation or application of (1.27) to (1.25) and (1.26) one gets finally the absolute translational and rotational acceleration of the rigid body  $K_i$  again written in the reference frame  $R$  as

$$\mathbf{a}_i^R(t) = \mathbf{a}_R^R + \left( \dot{\boldsymbol{\omega}}_{IR}^R + \tilde{\boldsymbol{\omega}}_{IR}^R \tilde{\boldsymbol{\omega}}_{IR}^R \right) \mathbf{r}_{Ri}^R + 2\tilde{\boldsymbol{\omega}}_{IR}^R \dot{\mathbf{r}}_{Ri}^R + \ddot{\mathbf{r}}_{Ri}^R, \quad (1.28)$$

$$\boldsymbol{\alpha}_{Ii}^R(t) = \dot{\boldsymbol{\omega}}_{IR}^R + \tilde{\boldsymbol{\omega}}_{IR}^R \boldsymbol{\omega}_{IR}^R + \dot{\boldsymbol{\omega}}_{Ri}^R. \quad (1.29)$$

Thus, in addition to the guidance and relative acceleration the Coriolis acceleration with the characteristic factor 2 is found for translations.

#### 1.4 Kinematics of Multibody Systems

So far only one free rigid body  $K_i$  was considered the position of which is uniquely described in the inertial frame  $I$  as

$$\mathbf{r}_i^I = [r_{i1} \ r_{i2} \ r_{i3}]^T, \quad \mathbf{S}^{Ii} \equiv \mathbf{S}_i = \mathbf{S}_i(\alpha_i, \beta_i, \gamma_i). \quad (1.30)$$

There are six position coordinates which are summarized in a  $6 \times 1$ - column matrix, simply called local position vector, as

$$\mathbf{x}_i(t) = [r_{i1} \ r_{i2} \ r_{i3} \ \alpha_i \ \beta_i \ \gamma_i]^T. \quad (1.31)$$

For a free multibody system consisting of  $p$  disassembled rigid bodies  $K_i$ ,  $i = 1(1)p$ , there remain  $6p$  position coordinates resulting in a  $6p \times 1$  global position vector of an unconstrained system

$$\mathbf{x}(t) = [\mathbf{x}_1^T \ \dots \ \mathbf{x}_p^T]^T. \quad (1.32)$$

Assembling the free system there appear constraints between the position coordinates and their derivatives. In realistic models of vehicles only holonomic constraints are found restricting the motion of the position coordinates by geometric or integrable kinematic constraints. These constraints are implicitly described by algebraical equations which may be time-dependent (rheonomic), too,

$$\varphi_j(\mathbf{x}, t) = 0, \quad j = 1(1)q. \quad (1.33)$$

Due to  $q$  constraints there remain  $f$  linear independent position coordinates characterizing  $f = 6p - q$  degrees of freedom. The  $f$  independent position coordinates are also called generalized coordinates and may be summarized in a  $f \times 1$ - column matrix as global position vector of the constraint system

$$\mathbf{y}(t) = [y_1 \ \dots \ y_f]^T. \quad (1.34)$$

By (1.33) and (1.34) the vector  $\mathbf{x}$  is an explicit function of the  $f$  generalized coordinates representing the constraints explicitly,

$$\mathbf{x} = \mathbf{x}(y, t) . \quad (1.35)$$

The choice of generalized coordinates is not unique. E.g., some of the local position coordinates (absolute coordinates) or differences between local coordinates (relative coordinates) may be chosen as generalized coordinates. However, there exists a unique relation between different sets of generalized coordinates represented by a regular, time-invariant  $f \times f$  -matrix  $\mathbf{T}$  resulting in the transformation

$$\mathbf{y}(t) = \mathbf{T} \bar{\mathbf{y}}(t) , \quad (1.36)$$

where  $\mathbf{y}$  and  $\bar{\mathbf{y}}$  are the corresponding global position vectors. The position variables (1.30) may be rewritten for the whole system as

$$\mathbf{r}_i(t) = \mathbf{r}_i(y, t) , \quad \mathbf{S}^{Ii}(t) \equiv \mathbf{S}_i(t) = \mathbf{S}_i(y, t) , \quad i = 1(1)p . \quad (1.37)$$

The corresponding velocity variables  $\{\mathbf{v}_i, \boldsymbol{\omega}_i\}$ ,  $\boldsymbol{\omega}_i = \boldsymbol{\omega}_{Ii}$  are obtained by differentiation as

$$\mathbf{v}_i(t) = \dot{\mathbf{r}}_i(t) = \frac{\partial \mathbf{r}_i}{\partial \mathbf{y}^T} \dot{\mathbf{y}} + \frac{\partial \mathbf{r}_i}{\partial t} = \mathbf{J}_{Ti}(\mathbf{y}, t) \dot{\mathbf{y}} + \bar{\mathbf{v}}_i(\mathbf{y}, t) , \quad (1.38)$$

$$\boldsymbol{\omega}_i(t) = \dot{\mathbf{s}}_i(t) = \frac{\partial \mathbf{s}_i}{\partial \mathbf{y}^T} \dot{\mathbf{y}} + \frac{\partial \mathbf{s}_i}{\partial t} = \mathbf{J}_{Ri}(\mathbf{y}, t) \dot{\mathbf{y}} + \bar{\boldsymbol{\omega}}_i(\mathbf{y}, t) , \quad (1.39)$$

where  $\partial \mathbf{s}_i$  describes the  $3 \times 1$ -vector of the infinitesimal rotation following from the rotation matrix analogously to the rotational velocity (1.17) as

$$\partial \bar{\mathbf{s}}_i = \partial \mathbf{S}_i \mathbf{S}_i^T := \begin{bmatrix} 0 & -\partial s_{i3} & \partial s_{i2} \\ \partial s_{i3} & 0 & -\partial s_{i1} \\ -\partial s_{i2} & \partial s_{i1} & 0 \end{bmatrix} , \quad \partial \mathbf{s}_i = \begin{bmatrix} \partial s_{i1} \\ \partial s_{i2} \\ \partial s_{i3} \end{bmatrix} . \quad (1.40)$$

The  $3 \times f$ -functional or Jacobian matrices  $\mathbf{J}_{Ti}$ ,  $\mathbf{J}_{Ri}$  of translation and rotation, respectively, identify the relation between the local and the generalized or global coordinates. The formation of these matrices is defined using the rules of matrix multiplication as shown for the translation matrix

$$\frac{\partial \mathbf{r}_i}{\partial \mathbf{y}^T} = \partial \mathbf{r}_i \left( \frac{1}{\partial \mathbf{y}^T} \right) = \mathbf{J}_{Ti} = \begin{bmatrix} \frac{\partial r_{i1}}{\partial y_1} & \frac{\partial r_{i1}}{\partial y_2} & \cdots & \frac{\partial r_{i1}}{\partial y_f} \\ \frac{\partial r_{i2}}{\partial y_1} & \frac{\partial r_{i2}}{\partial y_2} & \cdots & \frac{\partial r_{i2}}{\partial y_f} \\ \frac{\partial r_{i3}}{\partial y_1} & \frac{\partial r_{i3}}{\partial y_2} & \cdots & \frac{\partial r_{i3}}{\partial y_f} \end{bmatrix} . \quad (1.41)$$

From (1.38) and (1.39) one obtains by a second differentiation the acceleration variables  $\{\mathbf{a}_i, \boldsymbol{\alpha}_i\}$  depending on the position vector  $\mathbf{y}$  and its derivatives,

$$\mathbf{a}_i(t) = \dot{\mathbf{v}}_i(t) = \mathbf{J}_{Ti}(\mathbf{y}, t) \ddot{\mathbf{y}} + \frac{\partial \mathbf{v}_i}{\partial \mathbf{y}^T} \dot{\mathbf{y}} + \frac{\partial \mathbf{v}_i}{\partial t} , \quad (1.42)$$

$$\boldsymbol{\alpha}_i(t) = \dot{\boldsymbol{\omega}}_i(t) = \mathbf{J}_{Ri}(\mathbf{y}, t)\dot{\mathbf{y}} + \frac{\partial \boldsymbol{\omega}_i}{\partial \mathbf{y}^T} \dot{\mathbf{y}} + \frac{\partial \boldsymbol{\omega}_i}{\partial t}. \quad (1.43)$$

For scleronomic, time-invariant constraints the partial time derivatives in (1.38), (1.39), (1.42) and (1.43) are vanishing.

In addition to the real motions, the virtual motions are required in the next chapter dealing with dynamics. A virtual motion is defined as an arbitrary, infinitesimally small variation of the position completely compatible with the constraints at any time. Rheonomic constraints are considered to be frozen at the time under consideration. The symbol  $\delta$  of the virtual motion has the properties

$$\delta \mathbf{r} \neq \mathbf{0}, \quad \delta t \equiv 0. \quad (1.44)$$

The symbol  $\delta$  follows the rules of calculus, i.e., it yields

$$\delta(c\mathbf{r}) = c\delta\mathbf{r}, \quad \delta(\mathbf{r}_1 + \mathbf{r}_2) = \delta\mathbf{r}_1 + \delta\mathbf{r}_2, \quad \delta\mathbf{r}(\mathbf{y}) = \frac{\partial \mathbf{r}}{\partial \mathbf{y}^T} \delta\mathbf{y}. \quad (1.45)$$

Thus, the virtual motion of a multibody system reads as

$$\delta \mathbf{r}_i = \mathbf{J}_{Ti} \delta \mathbf{y}, \quad \delta \mathbf{s}_i = \mathbf{J}_{Ri} \delta \mathbf{y}, \quad i = 1(1)p. \quad (1.46)$$

This completes the kinematics for rigid body vehicle systems.

## 2 Dynamics

For the generation of the equations of motion of multibody systems, in addition to kinematics, the inertia of the bodies and the acting forces have to be considered. The Newton-Euler approach, also called the synthetic method, uses the free body diagram resulting in full set of local equations which may be reduced by the principles of d'Alembert and Jourdain to the equations of motion. The Lagrangian approach, representing the analytical method, is based on energy considerations and the equations of motions are found directly but without any information on the reaction forces.

### 2.1 Inertia Properties

The inertia of a rigid body  $K_i$  is characterized by its mass  $m_i$  and its inertia tensor  $\mathbf{I}_{Ci}$ . The coordinates of the inertia tensor read in the body-fixed frame  $\{C_i, \mathbf{e}_v^i\}$ , see Figure 4, as

$$\mathbf{I}_{Ci}^i = \int_{m_i} (\boldsymbol{\rho}^T \boldsymbol{\rho} \mathbf{E} - \boldsymbol{\rho} \boldsymbol{\rho}^T) dm = \begin{bmatrix} I_{11} & I_{12} & I_{13} \\ I_{21} & I_{22} & I_{23} \\ I_{31} & I_{32} & I_{33} \end{bmatrix}_{C_i} = \mathbf{const}. \quad (2.1)$$

The vector  $\boldsymbol{\rho} \equiv \boldsymbol{\rho}^i = [\rho_1 \ \rho_2 \ \rho_3]^T$  describes a material point with mass  $dm$  with respect to the center of mass  $C_i$  and  $\mathbf{E}$  means the  $3 \times 3$ -identity matrix. The inertia tensor  $\mathbf{I}_{Ci}^i$  is symmetric and positive definite, and constant in the body-fixed frame.

The coordinates of the inertia tensor depend on the mass distribution and on the choice of the reference frame. For a parallel displacement of the body-fixed frame from

the center of mass  $C_i$  to an arbitrary body-fixed point  $O_i$  characterized by the vector  $\mathbf{s}$  one gets

$$\mathbf{I}_{O_i}^i = \mathbf{I}_{C_i}^i + (\mathbf{s}^T \mathbf{s} \mathbf{E} - \mathbf{s} \mathbf{s}^T) m_i . \quad (2.2)$$

Thus, the diagonal elements of an inertia tensor are minimal for the center of mass.

For a homogeneous, purely rotational displacement by the rotation matrix  $\mathbf{S}^{ii'}$  from frame  $\mathbf{e}_v^i$  to  $\mathbf{e}_v^{i'}$  around the center of mass the transformation law for tensors applies as

$$\mathbf{I}_{C_i}^i = \mathbf{S}^{ii'} \mathbf{I}_{C_i}^{i'} \mathbf{S}^{i'i} \quad \text{or} \quad \mathbf{I}_{C_i}^{i'} = \mathbf{S}^{i'i} \mathbf{I}_{C_i}^i \mathbf{S}^{ii'} . \quad (2.3)$$

Please note that the inertia tensor may be time-variant if the frame  $\{C_i, \mathbf{e}_v^i\}$  is not body-fixed. This is especially true if the inertial frame is chosen,  $i' \equiv I$ , due to  $\mathbf{S}^{iI} = \mathbf{S}^{iI}(t)$ .

For all reference points there exists a special body-fixed frame in which the off-diagonal elements of the inertia tensor are vanishing, e.g.,

$$\mathbf{I}_{C_i} = \mathbf{diag} [I_1 \ I_2 \ I_3] = \mathbf{const} . \quad (2.4)$$

The remaining diagonal elements  $I_v$  are called principal moments of inertia with reference to  $C_i$  and the corresponding axes are the principal inertia axes. Both quantities follow from the eigenvalue problem

$$(I_v \mathbf{E} - \mathbf{I}_{C_i}^i) \mathbf{x}_v = \mathbf{0} . \quad (2.5)$$

Thus, the principal moments of inertia are the eigenvalues of the matrix  $\mathbf{I}_{C_i}^i$  and the eigenvectors  $\mathbf{x}_v = \mathbf{e}_v^i$  define the principal inertia axes which have to be unit vectors  $\mathbf{x}_v^T \mathbf{x}_v = 1$ .

## 2.2 Newton-Euler Equations

The synthetic method is based on the laws of Newton (1687) and Euler (1758) relating the translational motion represented by the momentum  $\mathbf{p}$  of a body  $K$  to the sum of the external forces  $\mathbf{f}$  and the rotational motion represented by the moment of momentum  $\mathbf{h}_O$  to the sum of the external torques  $\mathbf{l}_O$ ,

$$\frac{d^I}{dt} \mathbf{p} = \mathbf{f}, \quad \frac{d^I}{dt} \mathbf{h}_O = \mathbf{l}_O . \quad (2.6)$$

The time derivatives of the momentum  $\mathbf{p}$  and the moment of momentum  $\mathbf{h}_O$  have to be evaluated in the inertial frame  $I$ . The common reference point  $O$  of the moment of momentum and the resulting external torque may be an inertially fixed point like the origin of the inertial frame,  $O \equiv O_I$ , or the moving center of mass of the body,  $O \equiv C$ .

The fundamental laws (2.6) will now be applied to the rigid body  $K_i$ ,  $i = 1(1)p$ , of a multibody system and appropriate frames are chosen. First of all, the bodies  $K_i$  are dismantled and the constraints are replaced by reaction forces acting then externally on the bodies involved in the same amount but with opposite sign according to the counteraction principle (action = reaction). Further, the center of mass is used as reference point for all bodies,  $O \equiv C_i$ .

In the inertial frame  $I$  momentum and moment of momentum for a rigid body  $K_i$  using the inertia properties  $m_i, \mathbf{I}_{C_i}$  read as

$$\mathbf{p}_i^I = m_i \mathbf{v}_{C_i}^I, \quad m_i = \mathbf{const} , \quad (2.7)$$

$$\mathbf{h}_{C_i}^I = \mathbf{I}_{C_i}^I \boldsymbol{\omega}_i^I, \quad \mathbf{I}_{C_i}^I = \mathbf{I}_{C_i}^I(t). \quad (2.8)$$

where  $\mathbf{v}_{C_i}^I$  and  $\boldsymbol{\omega}_i^I$  mean absolute velocities. Introducing (2.7) and (2.8) in (2.6) and omitting the index  $C$  one finally gets Newton's and Euler's equations

$$m_i \dot{\mathbf{v}}_i^I = \mathbf{f}_i^I, \quad m_i = \text{const}, \quad (2.9)$$

$$\mathbf{I}_i^I \dot{\boldsymbol{\omega}}_i^I + \tilde{\boldsymbol{\omega}}_i^I \mathbf{I}_i^I \boldsymbol{\omega}_i^I = \mathbf{l}_i^I, \quad \mathbf{I}_i^I = \mathbf{I}_i^I(t). \quad (2.10)$$

In a second step these equations are transformed in a body-fixed frame resulting in

$$m_i \dot{\mathbf{v}}_i^i = \mathbf{f}_i^i, \quad m_i = \text{const}, \quad (2.11)$$

$$\mathbf{I}_i^i \dot{\boldsymbol{\omega}}_i^i + \tilde{\boldsymbol{\omega}}_i^i \mathbf{I}_i^i \boldsymbol{\omega}_i^i = \mathbf{l}_i^i, \quad \mathbf{I}_i^i = \mathbf{const}. \quad (2.12)$$

Equations (2.9) and (2.10), and (2.11) and (2.12) look completely identical. If there is only one body like in gyro dynamics, then, (2.12) is preferable due to the time-invariance of the inertia tensor. In multibody dynamics, however, this advantage is fading.

Equation (2.12) is also known as Euler's equation of gyro dynamics. It can be found with the moment of momentum given in the body-fixed frame from (2.6) directly using the law of differentiation in a rotating frame (1.27),

$$\frac{d^I}{dt} \mathbf{h}_i^i = \dot{\mathbf{h}}_i^i + \tilde{\boldsymbol{\omega}}_i^i \mathbf{h}_i^i = \mathbf{l}_i^i, \quad \mathbf{h}_i^i = \mathbf{I}_i^i \boldsymbol{\omega}_i^i. \quad (2.13)$$

Finally, in an arbitrarily moving reference frame  $R$  Newton's and Euler's equations are also available, see e.g. Schiehlen and Eberhard (2004),

$$m_i \ddot{\mathbf{r}}_{Ri} + m_i \left[ \mathbf{r}_{Ri}^{**} + \dot{\tilde{\boldsymbol{\omega}}}_R + \tilde{\boldsymbol{\omega}}_R \tilde{\boldsymbol{\omega}}_R \right] \mathbf{r}_{Ri} + 2\tilde{\boldsymbol{\omega}}_R \dot{\mathbf{r}}_{Ri} = \mathbf{f}_i, \quad (2.14)$$

$$\begin{aligned} & \mathbf{I}_i \dot{\boldsymbol{\omega}}_{Ri} + \tilde{\boldsymbol{\omega}}_{Ri} \mathbf{I}_i \boldsymbol{\omega}_{Ri} \\ & + [\mathbf{I}_i \boldsymbol{\omega}_R + \tilde{\boldsymbol{\omega}}_R \mathbf{I}_i \boldsymbol{\omega}_R + \tilde{\boldsymbol{\omega}}_R \boldsymbol{\omega}_{Ri} \text{sp} \mathbf{I}_i + 2\tilde{\boldsymbol{\omega}}_{Ri} \mathbf{I}_i \boldsymbol{\omega}_R] = \mathbf{l}_i \end{aligned} \quad (2.15)$$

Now, the coordinates of all vectors and tensors are related to the reference frame  $R$  where  $\mathbf{r}_{Ri}^{**}$  means that the second time derivation has to be made before in the inertial frame. As a matter of fact, a large number of additional inertia forces and torques appear due to the relative motion.

The Newton-Euler equations represent a set of  $6p$  scalar equations for  $6p$  unknowns which are composed of unknown velocity and position variables and unknown reaction forces and torques. In an unconstrained system reactions do not exist, i.e., there are  $6p$  ordinary differential equations (ODEs) to be solved. In a completely constrained system motion does not occur at all, i.e., altogether  $6p$  algebraical equations have to be solved. In vehicle dynamics, due to a certain number of constraints between the bodies, motions and reactions appear featuring a set of differential-algebraical equations (DAEs). However, by the principles of dynamics, a minimal set of  $f$  ODEs can be found facilitating the solution and simulation of the problem.

### 2.3 Principles of d'Alembert and Jourdain

Equations of motion represent a minimal set of ordinary differential equations (ODEs). They can be found from the Newton-Euler equations by elimination of the reaction forces and torques. This is achieved computationally efficient by the principles of dynamics considering the virtual work of a constrained multibody system. For this purpose the external forces acting on the dismantled bodies of the system are subdivided into applied forces  $\mathbf{f}_i^{(a)}$  and torques  $\mathbf{l}_i^{(a)}$  as well as reaction forces  $\mathbf{f}_i^{(r)}$  and torques  $\mathbf{l}_i^{(r)}$ . The latter ones do not contribute to the virtual work of the system

$$\delta W^r = \sum_{i=1}^p (\mathbf{f}_i^{(r)\top} \delta \mathbf{r}_i + \mathbf{l}_i^{(r)\top} \delta \mathbf{s}_i) = 0, \quad (2.16)$$

where the virtual motions  $\delta \mathbf{r}_i, \delta \mathbf{s}_i$  are known from (1.46). Equation (2.16) can be interpreted as a generalized orthogonality condition. For this purpose, in addition to the generalized coordinates  $\mathbf{y}$ , generalized reaction forces  $g_j, j = 1(1)q$ , are introduced and summarized in a  $q \times 1$ -vector as

$$\mathbf{g}(t) = [g_1 \dots \dots \dots g_q]^\top. \quad (2.17)$$

The number of generalized constraint forces is determined by the number  $q$  of constraints. The local constraint forces and torques follow from the implicit constraint equations (1.33) as

$$\mathbf{f}_i^{(r)\top} = \sum_{j=1}^q g_j \frac{\partial \varphi_j}{\partial \mathbf{r}_i^\top} = \sum_{j=1}^q g_j \frac{\partial \varphi_j}{\partial \mathbf{x}^\top} \frac{\partial \mathbf{x}}{\partial \mathbf{r}_i^\top} = \mathbf{g}^\top \mathbf{F}_{Ti}^\top, \quad (2.18)$$

$$\mathbf{l}_i^{(r)\top} = \sum_{j=1}^q g_j \frac{\partial \varphi_j}{\partial \mathbf{s}_i^\top} = \sum_{j=1}^q g_j \frac{\partial \varphi_j}{\partial \mathbf{x}^\top} \frac{\partial \mathbf{x}}{\partial \mathbf{s}_i^\top} = \mathbf{g}^\top \mathbf{F}_{Ri}^\top, \quad i = 1(1)p. \quad (2.19)$$

In matrix notation the  $3 \times q$ - Jacobians  $\mathbf{F}_{Ti}, \mathbf{F}_{Ri}$  are found from (2.18) and (2.19), and the condition (2.16) is rewritten as

$$\begin{aligned} \delta W^r &= \mathbf{g}^\top \sum_{i=1}^p (\mathbf{F}_{Ti}^\top \delta \mathbf{r}_i + \mathbf{F}_{Ri}^\top \delta \mathbf{s}_i) \\ &= \mathbf{g}^\top \sum_{i=1}^p (\mathbf{F}_{Ti}^\top \mathbf{J}_{Ti} + \mathbf{F}_{Ri}^\top \mathbf{J}_{Ri}) \delta \mathbf{y} = 0. \end{aligned} \quad (2.20)$$

Finally, the global  $6p \times q$ - distribution matrix  $\bar{\mathbf{Q}}$ , and the global  $6p \times f$ - Jacobian matrix  $\bar{\mathbf{J}}$  are introduced

$$\begin{aligned} \bar{\mathbf{Q}} &= [\mathbf{F}_{T1}^\top \dots \dots \mathbf{F}_{Tp}^\top \mathbf{F}_{R1}^\top \dots \dots \mathbf{F}_{Rp}^\top]^\top, \\ \bar{\mathbf{J}} &= [\mathbf{J}_{T1}^\top \dots \dots \mathbf{J}_{Tp}^\top \mathbf{J}_{R1}^\top \dots \dots \mathbf{J}_{Rp}^\top]^\top. \end{aligned} \quad (2.21)$$

Then, one gets from (2.20) simply

$$\bar{\mathbf{Q}}^\top \bar{\mathbf{J}} = \bar{\mathbf{J}}^\top \bar{\mathbf{Q}} = \mathbf{0}, \quad (2.22)$$



what clearly shows the generalized orthogonality between motion and constraint. The orthogonality condition or the vanishing virtual work, respectively, is independent from the coordinates chosen, and it is valid for all constrained mechanical systems.

D'Alembert's principle (1743) follows now from the Newton-Euler equations (2.9) and (2.10) after subdividing the external forces

$$\mathbf{f}_i = \mathbf{f}_i^{(e)} + \mathbf{f}_i^{(r)}, \quad \mathbf{l}_i = \mathbf{l}_i^{(e)} + \mathbf{l}_i^{(r)}, \quad (2.23)$$

and considering the orthogonality (2.16) as

$$\sum_{i=1}^p [(m_i \dot{\mathbf{v}}_i - \mathbf{f}_i^{(e)})^T \delta \mathbf{r}_i + (\mathbf{I}_i \dot{\boldsymbol{\omega}}_i + \tilde{\boldsymbol{\omega}}_i \mathbf{I}_i \boldsymbol{\omega}_i - \mathbf{l}_i^{(e)})^T \delta \mathbf{s}_i] = 0. \quad (2.24)$$

Obviously, the reaction forces are eliminated in (2.24).

Analogously Jourdain's principle (1908) can be stated which is based on the fact that the virtual power of the reaction forces is vanishing, too,

$$\delta P^r = \sum_{i=1}^p [\mathbf{f}_i^{(r)T} \delta' \mathbf{v}_i + \mathbf{l}_i^{(r)T} \delta' \boldsymbol{\omega}_i] = 0. \quad (2.25)$$

The virtual velocities  $\delta' \mathbf{v}_i, \delta' \boldsymbol{\omega}_i$  are arbitrary, infinitesimal small variations of the velocities completely compatible with the constraints at any time and at any position. Thus, it yields

$$\delta' \mathbf{v}_i \neq 0, \quad \delta' \boldsymbol{\omega}_i \neq 0, \quad \delta' \mathbf{r}_i \equiv 0, \quad \delta' \mathbf{s}_i \equiv 0, \quad \delta' t \equiv 0. \quad (2.26)$$

Moreover, the symbol  $\delta'$  follows the rules of calculus. Then, it remains Jourdain's principle as

$$\sum_{i=1}^p [(m_i \dot{\mathbf{v}}_i - \mathbf{f}_i^{(e)})^T \delta' \mathbf{v}_i + (\mathbf{I}_i \dot{\boldsymbol{\omega}}_i + \tilde{\boldsymbol{\omega}}_i \mathbf{I}_i \boldsymbol{\omega}_i - \mathbf{l}_i^{(e)})^T \delta' \boldsymbol{\omega}_i] = 0. \quad (2.27)$$

Similar to d'Alembert's principle all the reactions disappeared. However, the virtual displacements are replaced by the virtual velocities and the sometimes cumbersome evaluation of the virtual rotations is dropped. Further, Jourdain's principle handles nonlinear and nonholonomic constraints, too, which may appear in controlled vehicle systems.

In the American literature Jourdain's principle is referred to as Kane's equations and the virtual velocities are denoted as partial velocities, see Kane and Levinson (1985). Applying these principles for the generation of the equations of motion, the reactions have not to be considered at all. Therefore, the principles may be also classified as an analytical approach.

## 2.4 Energy Considerations and Lagrange's Equations

An alternative for the generation of the equations of motion is the analytical method by Lagrange (1788) based on energy considerations. The kinetic energy  $T$  of a rigid body reads as

$$T = \frac{1}{2} m v_C^2 + \frac{1}{2} \boldsymbol{\omega} \mathbf{I}_C \boldsymbol{\omega}, \quad (2.28)$$

where the inertia properties  $\{m, \mathbf{I}_C\}$  and the velocity properties  $\{\mathbf{v}_C, \boldsymbol{\omega}\}$  are related to the center of mass. The kinetic energy is composed by the translational and rotational energy of the body, it is a scalar quantity which may be computed in different frames, too.

The kinetic energy of a multibody system consisting of the bodies  $K_i, i = 1(1)p$ , comprises the kinetic energy of all bodies as

$$T = \frac{1}{2} \sum_{i=1}^p [(\mathbf{v}_i^I)^T m_i \mathbf{v}_i^I + (\boldsymbol{\omega}_i^I)^T \mathbf{I}_i^I \boldsymbol{\omega}_i^I], \quad (2.29)$$

written consistently in the inertial frame  $I$  and related to the center of mass  $C_i$  of each body  $K_i$ . If the work of the applied forces is independent of the path, then the forces have a potential  $U$  and it yields

$$\mathbf{f}^{(e)} = -\text{grad } U, \quad (2.30)$$

where  $U$  is a scalar function of the position. Forces satisfying (2.30) are called conservative, they do not change the total energy of the system. In contrary, non-conservative forces change the total energy, they are called dissipative since the total energy is decreasing. Conservative forces may be due to gravity,  $f_G = mg$ , or elasticity,  $f_F = -ks$ . The corresponding potentials read as

$$U_G = mgz, \quad U_F = \frac{1}{2}ks^2, \quad (2.31)$$

where  $z$  represents the vertical displacement of the center of mass of a body with mass  $m$  in the direction of gravity with acceleration  $g$ , and  $s$  means the displacement of an elastic spring with coefficient  $k$ . To the potentials a constant may be added, i.e., the origin of a potential can be arbitrarily chosen. The potential energy  $U$  of a multibody system is given by the sum of the body potentials  $U = \sum U_j$ . Multibody systems subject to conservative forces only are called conservative systems. For such systems it yields the energy conservation law

$$T + U = T_0 + U_0 = \text{const}. \quad (2.32)$$

The energy conservation law may be derived from Newton's and Euler's law, i.e., it does not contain no new information. Its application is advantageous for conservative systems with one degree of freedom to evaluate a relation between the position and velocity variable. If there are two different positions known, the unknown velocity can be found from (2.32),

Based on energy expressions, the equations of motion of multibody systems may be found, too. This will be shown for multibody systems with holonomic constraints. In contrary to the synthetic method, the body of the system have not to be dismantled, the system is considered as a whole. For this purpose the generalized coordinates  $\mathbf{y}$  are defined, and the position and the velocity variables (1.37), (1.38) and (1.39) are evaluated. As a result the kinetic energy is available as a function of  $y_k(t)$  and  $\dot{y}_k(t)$ ,  $k = 1(1)f$ ,

$$T = T(y_k, \dot{y}_k). \quad (2.33)$$

The applied forces and torques are projected in the direction of the generalized coordinates and composed to the generalized forces

$$q_k = \sum_{i=1}^p \left[ \left( \frac{\partial \mathbf{r}_i^I}{\partial y_k} \right)^T \mathbf{f}_i^{(e)I} + \left( \frac{\partial \mathbf{s}_i^I}{\partial y_k} \right)^T \mathbf{l}_i^{(e)I} \right], \quad k = 1(1)f, \quad (2.34)$$

where the rows of the Jacobian matrices (1.41) are used. The generalized forces may be also found by decomposition of the total work of the applied forces and torques

$$\delta W^e = \sum_{i=1}^p [\mathbf{f}_i^{(e)T} \delta \mathbf{r}_i + \mathbf{l}_i^{(e)T} \delta \mathbf{s}_i] = \sum_{k=1}^f q_k \delta y_k. \quad (2.35)$$

In any case the reaction forces and torques do not appear.

Now the Lagrangian equations of the second kind read as, see e.g. Magnus and Müller (1990),

$$\frac{d}{dt} \left( \frac{\partial T}{\partial \dot{y}_k} \right) - \frac{\partial T}{\partial y_k} = q_k, \quad k = 1(1)f. \quad (2.36)$$

For the evaluation of the equations of motion two partial and one total differentiations have to be performed with respect to one scalar function  $T(y_k, \dot{y}_k)$ . As a result the minimal number  $f$  of equations of motion is found. However, the reaction forces are completely lost and cannot be regained.

For conservative systems the generalized forces follow immediately from the potential

$$q_k = - \frac{\partial U}{\partial y_k}. \quad (2.37)$$

From (2.36) and (2.37) it remains

$$\frac{d}{dt} \left( \frac{\partial T}{\partial \dot{y}_k} \right) - \frac{\partial T}{\partial y_k} + \frac{\partial U}{\partial y_k} = 0, \quad k = 1(1)f. \quad (2.38)$$

Introducing the Lagrange function  $L = T - U$  also called the kinetic potential, then (2.38) is even more simplified and reads

$$\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{y}_k} \right) - \frac{\partial L}{\partial y_k} = 0, \quad L = T - U, \quad k = 1(1)f. \quad (2.39)$$

For some engineering applications it is advantageous to use surplus coordinates  $\bar{y}_j$ ,  $j = 1(1)f + r$ , in addition to the  $f$  generalized coordinates  $y_k$ ,  $k = 1(1)f$ . Then, there exist  $r$  geometric constraints between the surplus coordinates

$$\varphi_n = \varphi_n(\bar{y}_j) = 0, \quad n = 1(1)r. \quad (2.40)$$

The equations of motion are now extended by  $r$  Lagrangian multipliers  $\lambda_n$ ,  $n = 1(1)r$ , representing generalized constraint forces

$$\frac{d}{dt} \left( \frac{\partial T}{\partial \dot{\bar{y}}_j} \right) - \frac{\partial T}{\partial \bar{y}_j} = \bar{q}_j + \sum_{n=1}^r \lambda_n \frac{\partial \varphi_n}{\partial \bar{y}_j}, \quad j = 1(1)f + r. \quad (2.41)$$

Thus, a set of differential-algebraical equations remain.

### 3 Equations of Motion

In chapter 2 there has been presented two methods for the generation of the equations of motion, the synthetic method by Newton-Euler, and the analytic method by d'Alembert, Jourdain or Lagrange, respectively. The principal steps in the generation process by Newton-Euler and Lagrange are shown in Figure 9. Common starting point is the mechanical model of the vehicle composed by the elements of multibody systems. Common result are the equations of motion, they are identical with both methods if the same generalized coordinates are used. However, the effort is different. During the generation of the equations of motion using Lagrange's equations there appear terms in  $\frac{d}{dt} \left( \frac{\partial T}{\partial \dot{y}_i} \right)$  which are afterwards eliminated by  $\frac{\partial T}{\partial y_i}$  according to (2.36). This means an useless computational effort which is not required with the Newton-Euler approach, see e.g. Schiehlen and Eberhard (2004). On the other hand in the Newton-Euler equations the reactions have to be eliminated. Thus, both of these approaches have disadvantages which are avoided by a combination of the Newton-Euler equations with the principles presented in Section 2.3. The resulting equations of motion are always ordinary differential equations (ODEs). However, their form depends on the type of the multibody system. There are ideal and non-ideal systems, the first ones are characterized by applied forces and torques independent from any reaction while the second ones show a such dependency. E.g., gravitational forces, spring and damper forces are independent from any reactions while sliding friction forces and slip dependent contact forces, regularly found with tires in vehicle dynamics, are a function of the normal or reaction forces, respectively.

Within the class of ideal systems, ordinary and general multibody systems are distinguished. Ordinary multibody systems are due to holonomic constraints and applied forces depending only on position and velocity quantities, they can be always represented by a system of differential equations of the second order. For nonholonomic constraints and/or general force laws one gets general multibody systems.

The equations of ordinary multibody systems read as

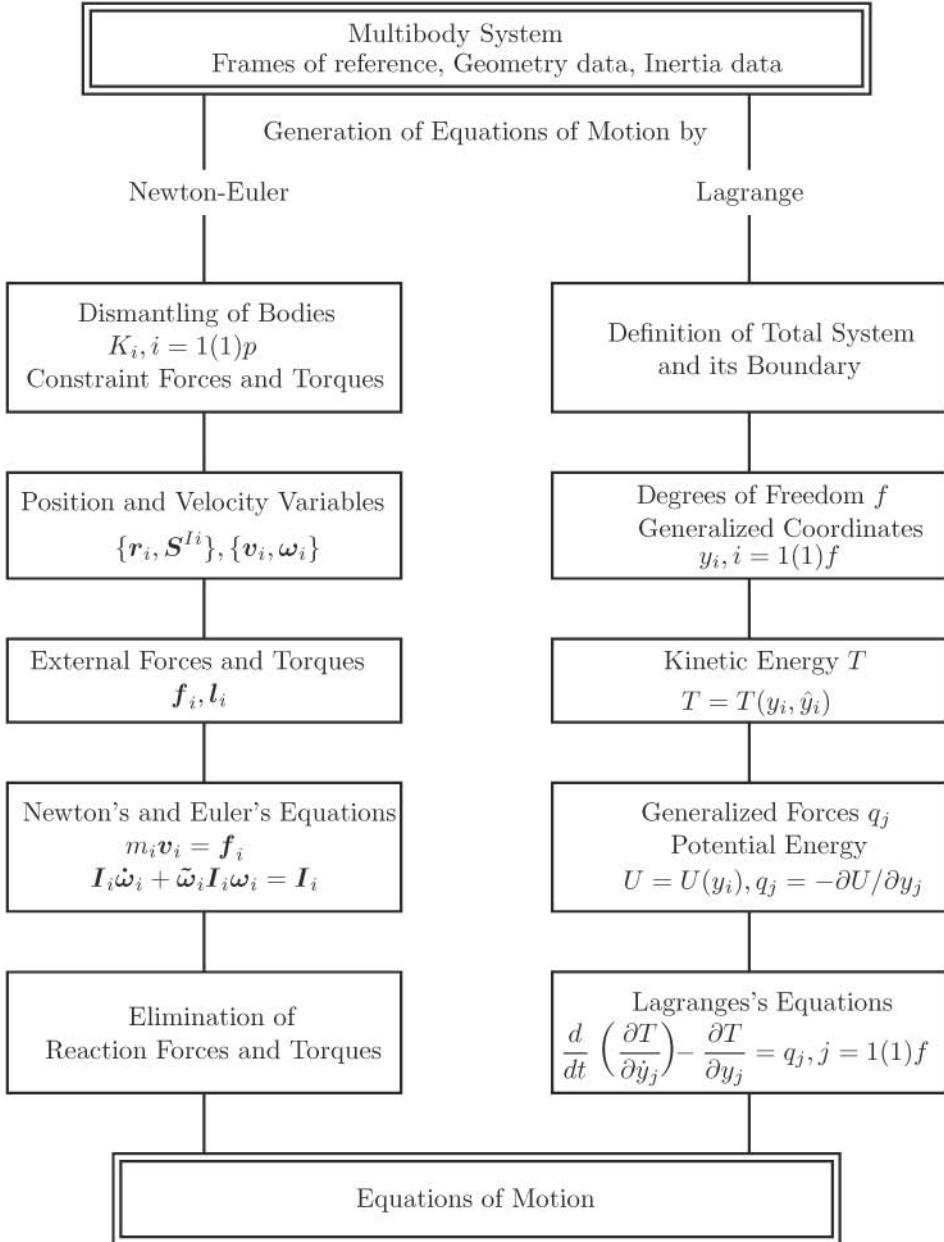
$$\mathbf{M}(\mathbf{y}, t) \ddot{\mathbf{y}}(t) + \mathbf{k}(\mathbf{y}, \dot{\mathbf{y}}, t) = \mathbf{q}(\mathbf{y}, \dot{\mathbf{y}}, t) , \quad (3.1)$$

where  $\mathbf{y}$  is the  $f \times 1$ -position vector of the generalized coordinates,  $\mathbf{M}$  is the  $f \times f$ -symmetric inertia matrix,  $\mathbf{k}$  is a  $f \times 1$ -vector of generalized gyroscopic forces including the Coriolis- and centrifugal forces as well as the gyroscopic torques, and the  $f \times 1$ -vector  $\mathbf{q}$  represents generalized applied forces. The equations of motion resulting from the analytical method have always the form (3.1) while the synthetic method often requires some calculations to get a symmetric inertia matrix.

In vehicle dynamics the deviations  $\tilde{\mathbf{y}}(t)$  from a reference motion  $\mathbf{y} = \mathbf{y}_R(t)$  are often small,

$$\mathbf{y}(t) = \mathbf{y}_R(t) + \tilde{\mathbf{y}}(t) . \quad (3.2)$$

Then, it follows by a Taylor series expansion under assumption of differentiable vector functions, and skipping of the second and higher order terms from (3.1) the linearized



**Figure 9.** Generation of equations of motion by the methods of Newton-Euler and Lagrange

equations of motion, see also Müller and Schiehlen (1985),

$$\mathbf{M}(t)\ddot{\tilde{\mathbf{y}}}(t) + \mathbf{P}(t)\dot{\tilde{\mathbf{y}}}(t) + \mathbf{Q}(t)\tilde{\mathbf{y}}(t) = \mathbf{h}(t), \quad (3.3)$$

where  $\mathbf{M}(t)$  is the symmetric, positive definite inertia matrix while  $\mathbf{P}(t)$  and  $\mathbf{Q}(t)$  characterize the velocity and position dependent forces and the vector  $\mathbf{h}(t)$  represents the external excitations. If all these matrices are time-invariant and subdivided in a symmetrical and skewsymmetrical part, then the equations of motion of a linear ordinary and time-invariant multibody system are found reading as

$$\mathbf{M}\ddot{\tilde{\mathbf{y}}}(t) + (\mathbf{D} + \mathbf{G})\dot{\tilde{\mathbf{y}}}(t) + (\mathbf{K} + \mathbf{N})\tilde{\mathbf{y}}(t) = \mathbf{h}(t), \quad (3.4)$$

where  $\tilde{\mathbf{y}}$  was simply replaced by  $\mathbf{y}$  and the  $f \times f$ -matrices have the properties

$$\mathbf{M} = \mathbf{M}^T > 0, \quad \mathbf{D} = \mathbf{D}^T, \quad \mathbf{G} = -\mathbf{G}^T, \quad \mathbf{K} = \mathbf{K}^T, \quad \mathbf{N} = -\mathbf{N}^T. \quad (3.5)$$

These matrices have a physical meaning which can be identified after premultiplication of (3.5) from the left by  $\dot{\mathbf{y}}^T$  resulting in the total time derivative of an energy expression

$$\underbrace{\dot{\mathbf{y}}^T \mathbf{M} \ddot{\mathbf{y}}}_{T} + \underbrace{\dot{\mathbf{y}}^T \mathbf{D} \dot{\mathbf{y}}}_{2R} + \underbrace{\dot{\mathbf{y}}^T \mathbf{G} \dot{\mathbf{y}}}_{0} + \underbrace{\dot{\mathbf{y}}^T \mathbf{K} \mathbf{y}}_{2S} + \underbrace{\dot{\mathbf{y}}^T \mathbf{N} \mathbf{y}}_{P} = \dot{\mathbf{y}}^T \mathbf{h}, \quad (3.6)$$

$$\frac{d}{dt}T + 2R + 0 + \frac{d}{dt}U + 2S = P. \quad (3.7)$$

The inertia matrix  $\mathbf{M}$  determines the kinetic energy  $T = \frac{1}{2}\dot{\mathbf{y}}^T \mathbf{M} \dot{\mathbf{y}}$  and therefore the inertia forces, from  $T > 0$  it follows again the positive definiteness of the inertia matrix. The damping matrix  $\mathbf{D}$  defines via Rayleigh's dissipation function  $R = \frac{1}{2}\dot{\mathbf{y}}^T \mathbf{D} \dot{\mathbf{y}}$  the damping forces while the gyro matrix  $\mathbf{G}$  describes the gyroscopic forces which do not change the total energy of the system. The stiffness matrix determines the potential energy  $U = \frac{1}{2}\mathbf{y}^T \mathbf{K} \mathbf{y}$  and, therefore, the conservative position forces while the matrix  $\mathbf{N}$  identifies the circulatory forces also known as nonconservative position forces. Furthermore,  $P$  describes the power of the external excitation forces. For  $\mathbf{D} = \mathbf{0}$ ,  $\mathbf{N} = \mathbf{0}$  and  $\mathbf{h} = \mathbf{0}$  the multibody system is conservative, i.e., the total energy is constant for all motions,

$$T + U = \text{const}. \quad (3.8)$$

The matrix properties (3.5) allow often to check the equations of motion with respect to the physical phenomena involved.

The equations of motion of ordinary multibody systems in nonlinear or linear form, Eqs. (3.1) or (3.3), respectively, are systems of differential equations of second order. For solution they have to be supplemented by the initial conditions for position and velocity,

$$\mathbf{y}(0) = \mathbf{y}_0, \quad \dot{\mathbf{y}}(0) = \dot{\mathbf{y}}_0 \quad (3.9)$$

The state vector of the vehicle is defined by the position vector and the velocity vector as

$$\mathbf{x}_F(t) = \begin{bmatrix} \mathbf{y}(t) \\ \dot{\mathbf{y}}(t) \end{bmatrix}, \quad (3.10)$$

where  $\mathbf{x}$  is the  $n \times 1$  -vector of the state variables. For ordinary multibody systems, and therefore for vehicles, too, it holds  $n = 2f$  where  $f$  means the number of degrees of freedom.

With the state vector (3.10) the equations of motion can be easily transferred into the corresponding state equations. In the nonlinear case it yields

$$\begin{aligned} \dot{\mathbf{y}}(t) &= \dot{\mathbf{y}}(t) \\ \underbrace{\ddot{\mathbf{y}}(t)} &= \underbrace{\mathbf{M}^{-1}(\mathbf{y}, t) [\mathbf{q}(\mathbf{y}, \dot{\mathbf{y}}, t) - \mathbf{k}(\mathbf{y}, \dot{\mathbf{y}}, t)]}_{\mathbf{a}_F(\mathbf{x}_F, t)}, \\ \dot{\mathbf{x}}_F(t) &= \end{aligned} \quad (3.11)$$

where a nonlinear  $n \times 1$ -vector function  $\mathbf{a}_F$  appears. For vehicles with small linear motions it remains

$$\dot{\mathbf{x}}_F(t) = \mathbf{A}_F \mathbf{x}_F(t) + \mathbf{B}_F \mathbf{u}_F(t). \quad (3.12)$$

where

$$\mathbf{A}_F = \begin{bmatrix} \mathbf{0} & \mathbf{E} \\ -\mathbf{M}^{-1}(\mathbf{K} + \mathbf{N}) & -\mathbf{M}^{-1}(\mathbf{D} + \mathbf{G}) \end{bmatrix} \quad (3.13)$$

is the  $n \times n$ -system matrix,  $\mathbf{B}_F$  the  $n \times r$ -input matrix and  $\mathbf{u}_F$  the  $r \times 1$ -input vector of the excitations acting on the vehicle.

For general multibody systems exist a broader variety of standard representations which will not be discussed in detail. General multibody systems can be uniquely represented by the state equations (3.11) or (3.12), too. However the special form of the system matrix  $A$  is no longer found.

## 4 Formalisms for Multibody Systems

The generation of equations of motion for large multibody systems is a nontrivial task requiring numerous steps during the evaluation of the fundamental relations. Beginning with the space age in the middle of the 1960s the generation of equations of motion was more formalized. The resulting formalisms were used for the development of computer codes for multibody systems, they are the basis of computational multibody dynamics. Twenty-five years later, in 1990, there were known 20 formalisms described in the Multibody System Handbook (Schiehlen, 1990). Many of them are used today.

Multibody system formalisms are based on Newton-Euler equations or Lagrange's equations, respectively, as described in Chapter 2 and 3. Regarding the computational procedure, numerical and symbolical formalisms are distinguished. Numerical formalisms supply the elements of the matrices as numbers in the case of linear time-invariant multibody systems (3.4). In the case of linear time-variant systems (3.3) and nonlinear systems (3.1) a numerical formalism provides the numbers in the equations of motion necessary for each time step required by the simulation programme. In contrary, symbolical formalisms generate the equations of motion only once with the computer how it is done with paper and pencil. The advantage is that variations of the system parameters and, for time-variant systems, the current time have to be inserted in the symbolical equations of motion only. Symbolical formalisms are especially helpful for optimizations and control design.

Furthermore, non-recursive and recursive formalisms are distinguished. Recursive formalisms make use of special topology properties of multibody systems.

#### 4.1 Non-recursive Formalisms

First of all the symbolical formalism NEWEUL is presented. NEWEUL is a research software based on the Newton-Euler equations and the principles of d'Alembert and Jourdain, it was developed at the University of Stuttgart, see e.g. Kreuzer and Leister (1991). NEWEUL generates equations of motion in minimal form (3.1) or (3.4), respectively, which may be solved by any integration code for ordinary differential equations. The formalisms comprises five steps which may be evaluated by hand for smaller multibody systems, too.

**Step 1:** System specification and input data. At first the multibody system is defined and treated as whole. The number of degrees of freedom is determined and the generalized coordinates  $y_k$ ,  $k = 1(1)f$ , are chosen. The inertial frame  $I$  and the body fixed frames  $i$  are chosen. Each body  $K_i$ ,  $i = 1(1)p$ , is dismantled, and the corresponding inertia parameters  $\{m_i, \mathbf{I}_i\}$ , the position variables  $\{\mathbf{r}_i^I, \mathbf{S}^{Ii}\}$  as well as the applied forces and torques  $\{\mathbf{f}_i^{(e)I}, \mathbf{l}_i^{(e)I}\}$  are specified. The quantities  $\mathbf{I}_i^I$  and  $\mathbf{l}_i^{(e)I}$  are related to the corresponding center of mass  $C_i$ . The input data read as

$$\mathbf{f}, \mathbf{y} = [y_1, \dots, y_j, \dots, y_f]^T, \quad (4.1)$$

$$p, \{m_i, \mathbf{I}_i^i\}, \{\mathbf{r}_i^I, \mathbf{S}^{Ii}\}, \{\mathbf{f}_i^{(e)I}, \mathbf{l}_i^{(e)I}\}, \quad i = 1(1)p, \quad (4.2)$$

$$\mathbf{r}_i^I = [r_{i1} \ r_{i2} \ r_{i3}]^T, \quad \mathbf{S}^{Ii} = \mathbf{S}^{Ii}(\alpha_i, \beta_i, \gamma_i) \equiv \mathbf{S}_i, \quad i = 1(1)p. \quad (4.3)$$

The rotation matrices are described by three angles, e.g. the Cardan angles  $\alpha_i, \beta_i, \gamma_i$ .

**Step 2:** Element consideration, local equations. At first the elements of the inertia tensors are computed in the inertial frame  $I$  by the transformation

$$\mathbf{I}_i^I = \mathbf{S}^{Ii} \mathbf{I}_i^i \mathbf{S}^{iI}. \quad (4.4)$$

Now there are all quantities available in the inertial frame  $I$  which is the only frame further used. Thus, the right upper index is no longer required and just skipped. The local equations of motion for each body  $K_i$  read as

$$m_i \dot{\mathbf{v}}_i = \mathbf{f}_i^{(e)} + \mathbf{f}_i^{(r)}, \quad i = 1(1)p, \quad (4.5)$$

$$\mathbf{I}_i \dot{\boldsymbol{\omega}}_i + \tilde{\boldsymbol{\omega}}_i \mathbf{I}_i = \mathbf{l}_i^{(e)} + \mathbf{l}_i^{(r)}, \quad i = 1(1)p, \quad (4.6)$$

where the external forces and torques are subdivided in the known applied forces and torques and the unknown reactions. The reactions are eliminated later and, therefore, they have not to be specified.

**Step 3:** Relation between local and global quantities. The relation between the position (4.3) of a single body  $K_i$  and the generalized coordinates (4.1) is given by the holonomic, and in general also rheonomic constraints. These relations are known from the input data as

$$\mathbf{r}_i = \mathbf{r}_i(\mathbf{y}, t), \quad \mathbf{S}_i = \mathbf{S}_i(\mathbf{y}, t), \quad i = 1(1)p, \quad (4.7)$$



and the corresponding velocities  $\{\mathbf{v}_i, \boldsymbol{\omega}_i\}$  and accelerations  $\{\mathbf{a}_i, \boldsymbol{\alpha}_i\}$  are computed as

$$\mathbf{v}_i = \dot{\mathbf{r}}_i = \frac{\partial \mathbf{r}_i}{\partial \mathbf{y}^\top} \dot{\mathbf{y}} + \frac{\partial \mathbf{r}_i}{\partial t} = \mathbf{J}_{T_i}(\mathbf{y}, t) \dot{\mathbf{y}} + \bar{\mathbf{v}}_i(\mathbf{y}, t), \quad (4.8)$$

$$\boldsymbol{\omega}_i = \dot{\mathbf{s}}_i = \frac{\partial \mathbf{s}_i}{\partial \mathbf{y}^\top} \dot{\mathbf{y}} + \frac{\partial \mathbf{s}_i}{\partial t} = \mathbf{J}_{R_i}(\mathbf{y}, t) \dot{\mathbf{y}} + \bar{\boldsymbol{\omega}}_i(\mathbf{y}, t), \quad (4.9)$$

$$\mathbf{a}_i = \dot{\mathbf{v}}_i = \mathbf{J}_{T_i}(\mathbf{y}, t) \ddot{\mathbf{y}} + \frac{\partial \mathbf{v}_i}{\partial \mathbf{y}^\top} \dot{\mathbf{y}} + \frac{\partial \mathbf{v}_i}{\partial t}, \quad (4.10)$$

$$\boldsymbol{\alpha}_i = \dot{\boldsymbol{\omega}}_i = \mathbf{J}_{R_i}(\mathbf{y}, t) \ddot{\mathbf{y}} + \frac{\partial \boldsymbol{\omega}_i}{\partial \mathbf{y}^\top} \dot{\mathbf{y}} + \frac{\partial \boldsymbol{\omega}_i}{\partial t}. \quad (4.11)$$

For scleronomic constraints the partial time-derivatives are disappearing.

The  $3 \times f$ -Jacobian matrices  $\mathbf{J}_{T_i}, \mathbf{J}_{R_i}$  of translation and rotation, respectively, present the relation between the local and global coordinates as shown by the virtual motion

$$\delta \mathbf{r}_i = \mathbf{J}_{T_i} \delta \mathbf{y}, \quad \delta \mathbf{s}_i = \mathbf{J}_{R_i} \delta \mathbf{y}, \quad i = 1(1)p. \quad (4.12)$$

These matrices are available from kinematics by relations (4.8) and (4.9). After these preparatory computations the local equations (4.5) and (4.6) of body  $K_i$  of the multibody system are evaluated as functions of the generalized coordinates and their derivatives.

**Step 4:** System consideration, global equations. At first the local equations, depending on the generalized coordinates, are composed in global matrices and vectors. For this purpose the  $6p \times 6p$ -diagonal matrix  $\overline{\overline{\mathbf{M}}}$  of the inertia quantities is introduced

$$\overline{\overline{\mathbf{M}}} = \text{diag}[m_1 \mathbf{E}, m_2 \mathbf{E}, \dots, m_p \mathbf{E}, \mathbf{I}_1, \mathbf{I}_2, \dots, \mathbf{I}_p], \quad (4.13)$$

where  $\mathbf{E}$  means the  $3 \times 3$ -unit matrix. Further, the  $6p \times 1$ -vectors  $\bar{\mathbf{q}}^{(e)}, \bar{\mathbf{q}}^{(r)}, \bar{\mathbf{k}}$  are used to summarize all applied, reaction, Coriolis and gyro forces and torques. These three vectors are defined as follows

$$\bar{\mathbf{q}} = [\mathbf{f}_1^\top, \dots, \mathbf{f}_p^\top, \mathbf{l}_1^\top, \dots, \mathbf{l}_p^\top]^\top. \quad (4.14)$$

And finally the global  $6p \times f$ -Jacobian matrix is introduced as

$$\bar{\mathbf{J}} = [\mathbf{J}_{T_1}^\top, \dots, \mathbf{J}_{T_p}^\top, \mathbf{J}_{R_1}^\top, \dots, \mathbf{J}_{R_p}^\top]^\top. \quad (4.15)$$

The global Newton-Euler equations are now represented as one  $6p \times 1$ -vector equation

$$\overline{\overline{\mathbf{M}}} \bar{\mathbf{J}} \ddot{\mathbf{y}} + \bar{\mathbf{k}}(\mathbf{y}, \dot{\mathbf{y}}, t) = \bar{\mathbf{q}}^{(e)}(\mathbf{y}, \dot{\mathbf{y}}, t) + \bar{\mathbf{q}}^{(r)}. \quad (4.16)$$

These  $6p$  equations are reduced to the minimal number of  $f$  ordinary differential equations by left premultiplication with the transposed  $f \times 6p$ -Jacobian matrix  $\bar{\mathbf{J}}^\top$ ,

$$\bar{\mathbf{J}}^\top \overline{\overline{\mathbf{M}}} \bar{\mathbf{J}} \ddot{\mathbf{y}} + \bar{\mathbf{J}}^\top \bar{\mathbf{k}}(\mathbf{y}, \dot{\mathbf{y}}, t) = \bar{\mathbf{J}}^\top \bar{\mathbf{q}}^{(e)}(\mathbf{y}, \dot{\mathbf{y}}, t), \quad (4.17)$$

where the term  $\bar{\mathbf{J}}^\top \bar{\mathbf{q}}^{(r)}$  is vanishing due to (2.16).

Summarizing the matrix products, one gets the equations of motion with symmetric inertia matrix

$$\begin{aligned} \mathbf{M}(\mathbf{y}, t)\ddot{\mathbf{y}}(t) + \mathbf{k}(\mathbf{y}, \dot{\mathbf{y}}, t) &= \mathbf{q}(\mathbf{y}, \dot{\mathbf{y}}, t), \\ \mathbf{M} = \mathbf{M}^T = \overline{\mathbf{J}}^T \overline{\overline{\mathbf{M}}} \overline{\mathbf{J}}, \quad \mathbf{k} &= \overline{\mathbf{J}}^T \overline{\mathbf{k}}, \quad \mathbf{q} = \overline{\mathbf{J}}^T \overline{\mathbf{q}}^{(e)}, \end{aligned} \quad (4.18)$$

in complete agreement with (3.1).

**Optional step 5:** Computation of the reaction forces and torques. In contrary to Lagrange's approach, the reactions can be regained if the global distribution matrix  $\overline{\mathbf{Q}}$  according to (2.16) and the generalized constraint force vector  $\mathbf{g}$  from (2.17) are employed. Then, it yields

$$\overline{\mathbf{q}}^{(r)} = \overline{\mathbf{Q}}\mathbf{g}. \quad (4.19)$$

Due to the orthogonality condition (2.22) the left premultiplication of (4.16) with  $\overline{\mathbf{Q}}^T \overline{\overline{\mathbf{M}}}^{-1}$  results in linear, completely algebraic reaction equations

$$\begin{aligned} \hat{\mathbf{N}}(\mathbf{y}, t)\mathbf{g}(t) + \hat{\mathbf{q}}(\mathbf{y}, \dot{\mathbf{y}}, t) &= \hat{\mathbf{k}}(\mathbf{y}, \dot{\mathbf{y}}, t), \\ \hat{\mathbf{N}} = \hat{\mathbf{N}}^T = \overline{\mathbf{Q}}^T \overline{\overline{\mathbf{M}}}^{-1} \overline{\mathbf{Q}}, \quad \hat{\mathbf{q}} &= \overline{\mathbf{Q}}^T \overline{\overline{\mathbf{M}}}^{-1} \overline{\mathbf{q}}^{(e)}, \quad \hat{\mathbf{k}} = \overline{\mathbf{Q}}^T \overline{\overline{\mathbf{M}}}^{-1} \overline{\mathbf{k}}, \end{aligned} \quad (4.20)$$

where  $\hat{\mathbf{N}}$  is the symmetric, generally positive definite  $q \times q$ -reaction matrix while the  $q \times 1$ -vectors  $\hat{\mathbf{q}}$ ,  $\hat{\mathbf{k}}$  show the influence of the applied and gyroscopic forces and torques on the reactions. On the other hand, by elimination of  $\ddot{\mathbf{y}}$  in (4.16) one gets

$$\overline{\mathbf{q}}^{(r)} = \overline{\mathbf{Q}}\mathbf{g} = [\mathbf{E} - \overline{\overline{\mathbf{M}}} \overline{\mathbf{J}} (\overline{\mathbf{J}}^T \overline{\overline{\mathbf{M}}} \overline{\mathbf{J}})^{-1} \overline{\mathbf{J}}^T] (\overline{\mathbf{k}} - \overline{\mathbf{q}}^{(e)}). \quad (4.21)$$

Then, by partitioning of (4.21) only some of the constraint forces and torques may be computed.

In vehicle dynamics often contact forces are found resulting in nonideal multibody systems. Then, the  $f$  equations of motion (4.18) have to be solved simultaneously with the  $q$  equations of reaction (4.20).

The presented five steps show that the NEWEUL formalism is based on the Newton-Euler equations, however, they are supplemented by typical features of the analytical method like generalized coordinates and generalized reactions. The required computations include summation, multiplication and differentiation of vectors and matrices, simplification of trigonometrical expressions and linearization of expressions. This computations are symbolically performed in NEWEUL. But the underlying formalism may be also executed by any formula manipulation software, e.g. by MAPLE. More details on NEWEUL are found on the web: [www.itm.uni-stuttgart.de/research/neweul](http://www.itm.uni-stuttgart.de/research/neweul).

The most widely used numerical formalism in vehicle engineering is MSC.ADAMS. The software MSC.ADAMS generate numerical solutions of the original non-reduced Newton-Euler equations (2.9) and (2.10) rewritten as

$$\overline{\overline{\mathbf{M}}} \ddot{\mathbf{x}} + \overline{\mathbf{k}}(\mathbf{x}, \dot{\mathbf{x}}, t) = \overline{\mathbf{q}}^{(e)}(\mathbf{x}, \dot{\mathbf{x}}, t) + \overline{\mathbf{Q}}\mathbf{g}, \quad \overline{\mathbf{Q}}^T = -\frac{\partial \overline{\Phi}}{\partial \mathbf{x}^T} = -\overline{\Phi}_{\mathbf{x}} \quad (4.22)$$

where the  $6p \times 1$ -vector  $\mathbf{x}(t)$  is composed by Cartesian coordinates and the  $6p \times q$ -distribution matrix  $\overline{\mathbf{Q}}$  of the reactions is found from the implicit constraints (1.33). After two total time derivatives of (1.33) a set of  $6p + q$  linear equations remains for the unknowns  $\ddot{\mathbf{x}}$  and  $\mathbf{g}$ ,

$$\begin{bmatrix} \overline{\mathbf{M}} & \Phi_x^T \\ \Phi_x & \mathbf{0} \end{bmatrix} \begin{bmatrix} \ddot{\mathbf{x}} \\ \mathbf{g} \end{bmatrix} = \begin{bmatrix} \overline{\mathbf{q}}^{(e)} - \overline{\mathbf{k}} \\ -\Phi_t - \dot{\Phi}_x \dot{\mathbf{x}} \end{bmatrix}. \quad (4.23)$$

Here,  $\overline{\mathbf{M}}$  is a blockdiagonal  $6p \times 6p$ -matrix which allows the application of sparse matrix techniques. Eqs. (4.23) represent differential-algebraical equations which are solved within the MSC.ADAMS software with special integration codes. As with all commercial codes a graphical interface is used, and mathematical equations are not to be handled. For more details search for ADAMS under: [www.mscsoftware.com](http://www.mscsoftware.com)

## 4.2 Recursive Formalisms

For time integration of holonomic systems the inertia matrix in (3.1) or (4.18), respectively, has to be inverted what is numerically costly for systems with many degrees of freedom,

$$\ddot{\mathbf{y}}(t) = \mathbf{M}^{-1}(\mathbf{y}, t) [\mathbf{q}(\mathbf{y}, \dot{\mathbf{y}}, t) - \mathbf{k}(\mathbf{y}, \dot{\mathbf{y}}, t)]. \quad (4.24)$$

Recursive algorithms avoid this matrix inversion. The fundamental requirement, however, is a chain or tree topology of the multibody system as shown in Figure 10. Loop topologies are not included. Contributions on recursive algorithms are due, e.g., to Hollerbach (1980), Bae and Haug (1987a), Brandl et al. (1988), Schiehlen (1991).

**Kinematics** Recursive kinematics uses the relative motion between two neighboring bodies and the related constraints as shown in Figure 11. The absolute translational and rotational velocity vector  $\mathbf{w}_i$  of body  $i$ , also denoted as twist, is related to the absolute velocity vector  $\mathbf{w}_{i-1}$  of body  $i-1$  and the generalized coordinates  $\mathbf{y}_i$  of the joint  $i$  between this two bodies. It yields

$$\underbrace{\begin{bmatrix} \mathbf{v}_{Oi} \\ \boldsymbol{\omega}_i \end{bmatrix}}_{\mathbf{w}_i} = \underbrace{\mathbf{S}^{i,i-1} \begin{bmatrix} \mathbf{E} & -\tilde{\mathbf{r}}_{Oi-1,Oi} \\ \mathbf{0} & \mathbf{E} \end{bmatrix}}_{\mathbf{C}_i} \underbrace{\begin{bmatrix} \mathbf{v}_{Oi-1} \\ \boldsymbol{\omega}_{i-1} \end{bmatrix}}_{\mathbf{w}_{i-1}} + \underbrace{\mathbf{S}^{i,i-1} \begin{bmatrix} \mathbf{J}_{Ti} \\ \mathbf{J}_{Ri} \end{bmatrix}}_{\mathbf{J}_i} \dot{\mathbf{y}}_i \quad (4.25)$$

Using the fundamentals of relative motion of rigid bodies, it remains for the absolute acceleration

$$\mathbf{b}_i = \mathbf{C}_i \mathbf{b}_{i-1} + \mathbf{J}_i \ddot{\mathbf{y}}_i + \boldsymbol{\beta}_i(\dot{\mathbf{y}}_i, \mathbf{w}_{i-1}), \quad i = 1(1)p \quad (4.26)$$

where the vector  $\mathbf{b}_i$  summarizes the translational and rotational accelerations of body  $i$  as well.

For the total system one gets for the absolute acceleration in matrix notation

$$\mathbf{b} = \mathbf{C} \mathbf{b} + \mathbf{J} \ddot{\mathbf{y}} + \boldsymbol{\beta} \quad (4.27)$$

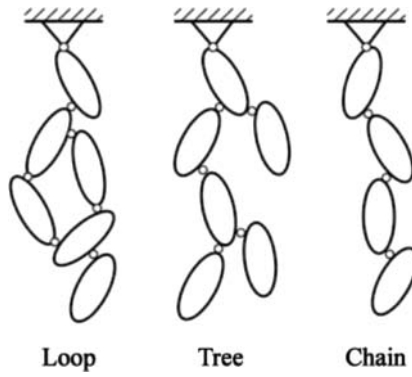


Figure 10. Topologies of multibody systems

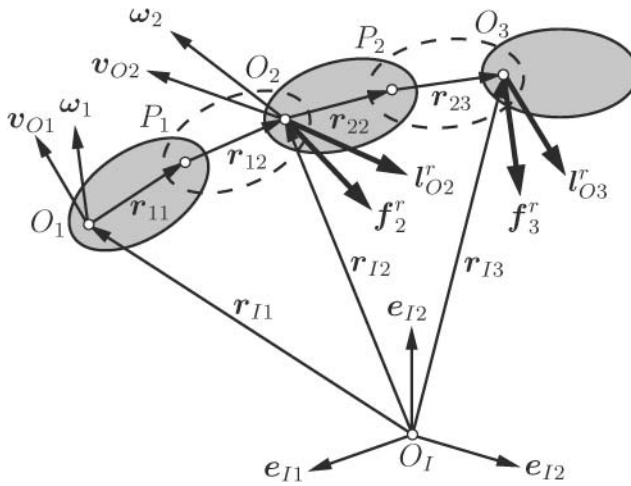


Figure 11. Absolute twists and constraint wrenches for neighboring bodies

where the geometry matrix  $\mathbf{C}$  is a lower block-sub-diagonal matrix and the Jacobian  $\mathbf{J}$  is a block-diagonal matrix as follows

$$\mathbf{C} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{C}_2 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_3 & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{C}_p & \mathbf{0} \end{bmatrix}, \quad \mathbf{J} = \begin{bmatrix} \mathbf{J}_1 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_2 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{J}_3 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{J}_p \end{bmatrix}. \quad (4.28)$$

From (4.27) it follows the non-recursive form of the absolute accelerations as

$$\mathbf{b} = (\mathbf{E} - \mathbf{C})^{-1} \mathbf{J} \ddot{\mathbf{y}} + \bar{\boldsymbol{\beta}} \quad (4.29)$$

where the global Jacobian matrix  $\bar{\mathbf{J}}$  is found again, see (2.21) or (4.15)

$$\bar{\mathbf{J}} = (\mathbf{E} - \mathbf{C})^{-1} \mathbf{J} = \begin{bmatrix} \mathbf{J}_1 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{C}_2 \mathbf{J}_1 & \mathbf{J}_2 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{C}_3 \mathbf{C}_2 \mathbf{J}_1 & \mathbf{C}_3 \mathbf{J}_2 & \mathbf{J}_3 & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ * & * & * & \cdots & \mathbf{J}_p \end{bmatrix}. \quad (4.30)$$

Due to the chain topology the global Jacobian matrix is a lower triangular matrix.

**Newton-Euler Equations** Newton's and Euler's equations are now written for body  $i$  in its body fixed frame at the joint position  $O_i$  using the absolute accelerations and the external forces and torques summarized in the vector  $\mathbf{q}_i$ , denoted also as wrench, acting on the body with holonomic constraints:

$$\underbrace{\begin{bmatrix} m_i \mathbf{E} & m_i \tilde{\mathbf{r}}_{O_i C_i}^T \\ m_i \tilde{\mathbf{r}}_{O_i C_i} & \mathbf{I}_{O_i} \end{bmatrix}}_{\mathbf{M}_i = \text{const}} \underbrace{\begin{bmatrix} \mathbf{a}_{O_i} \\ \boldsymbol{\alpha}_i \end{bmatrix}}_{\mathbf{b}_i} + \underbrace{\begin{bmatrix} m_i \tilde{\boldsymbol{\omega}}_i \tilde{\boldsymbol{\omega}}_i \mathbf{r}_{O_i C_i} \\ \tilde{\boldsymbol{\omega}}_i \mathbf{I}_{O_i} \boldsymbol{\omega}_i \end{bmatrix}}_{\mathbf{k}_i} = \underbrace{\begin{bmatrix} \mathbf{f}_i \\ \mathbf{l}_{O_i} \end{bmatrix}}_{\mathbf{q}_i}. \quad (4.31)$$

Moreover, the external forces are composed of applied forces  $\mathbf{q}_i^{(e)}$  and constraints forces  $\mathbf{q}_i^{(r)}$  where the generalized constraint forces of the joint  $i$  and joint  $i - 1$  appear:

$$\mathbf{q}_i = \mathbf{q}_i^{(e)} + \mathbf{q}_i^{(r)}, \quad \mathbf{q}_i^{(r)} = \mathbf{Q}_i \mathbf{g}_i - \mathbf{C}_{i+1}^T \mathbf{Q}_{i+1} \mathbf{g}_{i+1}, \quad i = 1(1)p. \quad (4.32)$$

**Equations of Motion** For the total system a set of  $18p$  scalar equations remains from (4.27), (4.31) and (4.32)

$$\mathbf{b} = \bar{\mathbf{J}} \ddot{\mathbf{y}} + \bar{\boldsymbol{\beta}}, \quad (4.33)$$

$$\bar{\bar{\mathbf{M}}} \mathbf{b} + \bar{\mathbf{k}} = \mathbf{q}^{(e)} + \mathbf{q}^{(r)}, \quad (4.34)$$

$$\mathbf{q}^{(r)} = (\mathbf{E} - \mathbf{C})^T \mathbf{Q} \mathbf{g} = \bar{\mathbf{Q}} \mathbf{g}. \quad (4.35)$$

with  $18p$  unknowns in the vectors  $\mathbf{b}$ ,  $\mathbf{y}$ ,  $\mathbf{q}^{(r)}$ ,  $\mathbf{g}$ .

Now (4.33) and (4.35) are inserted in (4.34) and the global orthogonality  $\bar{\mathbf{J}}^T \bar{\mathbf{Q}} = \mathbf{0}$  is used again resulting in the standard form (4.18) of the equations of motion. The mass matrix is completely full, again, and the vector  $\mathbf{k}$  depends not only on the generalized velocities but also on the absolute velocities,

$\mathbf{M} =$

$$\begin{bmatrix} \mathbf{J}_1^T (\mathbf{M}_1 + \mathbf{C}_2^T (\mathbf{M}_2 + \mathbf{C}_3^T \mathbf{M}_3 \mathbf{C}_3) \mathbf{C}_2) \mathbf{J}_1 & \mathbf{J}_1^T \mathbf{C}_2^T (\mathbf{M}_2 + \mathbf{C}_3^T \mathbf{M}_3 \mathbf{C}_3) \mathbf{J}_2 & \mathbf{J}_1^T \mathbf{C}_2^T \mathbf{C}_3^T \mathbf{M}_3 \mathbf{J}_3 \\ \mathbf{J}_2^T (\mathbf{M}_2 + \mathbf{C}_3^T \mathbf{M}_3 \mathbf{C}_3) \mathbf{C}_2 \mathbf{J}_1 & \mathbf{J}_1^T (\mathbf{M}_2 + \mathbf{C}_3^T \mathbf{M}_3 \mathbf{C}_3) \mathbf{J}_2 & \mathbf{J}_2^T \mathbf{C}_3^T \mathbf{M}_3 \mathbf{J}_3 \\ \mathbf{J}_3^T \mathbf{M}_3 \mathbf{C}_3 \mathbf{C}_2 \mathbf{J}_1 & \mathbf{J}_3^T \mathbf{M}_3 \mathbf{C}_3 \mathbf{J}_2 & \mathbf{J}_3^T \mathbf{M}_3 \mathbf{J}_3 \end{bmatrix}, \quad (4.36)$$

$$\mathbf{k} = \mathbf{k}(\mathbf{y}, \dot{\mathbf{y}}, \mathbf{w}) . \quad (4.37)$$

However, the mass matrix shows now a characteristic structure which can be used for a Gauss transformation.

**Recursion** There are three steps required to obtain the generalized accelerations.

1. Forward recursion to get the absolute motion starting with  $i = 1$ .
2. Backward recursion using a Gauss transformation starting with  $i = p$ . As a result the system

$$\hat{\mathbf{M}} \ddot{\mathbf{y}} + \hat{\mathbf{k}} = \hat{\mathbf{q}} \quad (4.38)$$

is obtained where  $\hat{\mathbf{M}}$  is a lower triangular matrix

$$\hat{\mathbf{M}} = \begin{bmatrix} \mathbf{J}_1^T \tilde{\mathbf{M}}_1 \mathbf{J}_1 & 0 & 0 \\ \mathbf{J}_2^T \tilde{\mathbf{M}}_2 \mathbf{C}_2 \mathbf{J}_1 & \mathbf{J}_2^T \tilde{\mathbf{M}}_2 \mathbf{J}_2 & 0 \\ \mathbf{J}_3^T \tilde{\mathbf{M}}_3 \mathbf{C}_3 \mathbf{C}_2 \mathbf{J}_1 & \mathbf{J}_3^T \tilde{\mathbf{M}}_3 \mathbf{C}_3 \mathbf{J}_2 & \mathbf{J}_3^T \tilde{\mathbf{M}}_3 \mathbf{J}_3 \end{bmatrix}, \quad (4.39)$$

the block elements of which follow from the recursion formula.

$$\tilde{\mathbf{M}}_{i-1} = \mathbf{M}_{i-1} + \mathbf{C}_i^T (\tilde{\mathbf{M}}_i - \tilde{\mathbf{M}}_i \mathbf{J}_i (\mathbf{J}_i^T \tilde{\mathbf{M}}_i \mathbf{J}_i)^{-1} \mathbf{J}_i^T \tilde{\mathbf{M}}_i) \mathbf{C}_i, \quad (4.40)$$

3. Forward recursion for the generalized accelerations starting with  $i = 1$ . The recursion requires some computational overhead. Therefore, the recursive algorithms are more efficient than the matrix inversion for more than  $p = 8 - 10$  bodies.

There are also some extensions of the recursive approach to loop topologies, see Bae and Haug (1987b) and Saha and Schiehlen (2001). Further, it has to be mentioned that there are two commercial codes based on recursive formalisms: SIMPACK and RecurDyn. For more details see: [www.simpack.de](http://www.simpack.de) and [www.functionbay.co.kr](http://www.functionbay.co.kr)

## 5 Modelling of Random Guideways

Vehicles running on guideways are subject to different types of disturbances. One can distinguish single events like curbs, dips or holes that are described mathematically as step functions or impulse functions, they excite primarily free vibrations of the vehicle. By

far more important are the permanent acting guideway irregularities, they cause forced excitations of the vehicle. In the early days they have been assumed to be sinusoidal and were investigated in a deterministic way. Today they are described as random processes and they are analyzed by stochastic methods. Although the guideway irregularities can locally be measured in an exact and reproducible way and, thus, they can also be described deterministically. However, of interest is not the local but the global behaviour. Here, random processes are the adequate mathematical description. In the following the guideway irregularities will at first be described as a roughness profile depending on a spacial coordinate. Then it will be transformed into the time domain regarding the vehicle speed and resulting in a random excitation process acting on the vehicle and generating random vibrations. The resulting system response is important to assess e. g. ride comfort and ride safety.

The investigation of road roughness profiles is going on for decades, see Mitschke (1972), Braun (1969), Voy (1978) and Bormann (1978). Numerous measurements have shown that a mathematical description can be given in form of

- stationary, normal distributed and ergodic random processes.

Before the roughness profiles and the corresponding excitation processes are described a general introduction into random processes shall be given in order to explain the fore-mentioned properties, see also Crandall and Mark (1963), Newland (1975) and Heinrich and Hennig (1978).

### 5.1 Mathematical Description of Random Processes

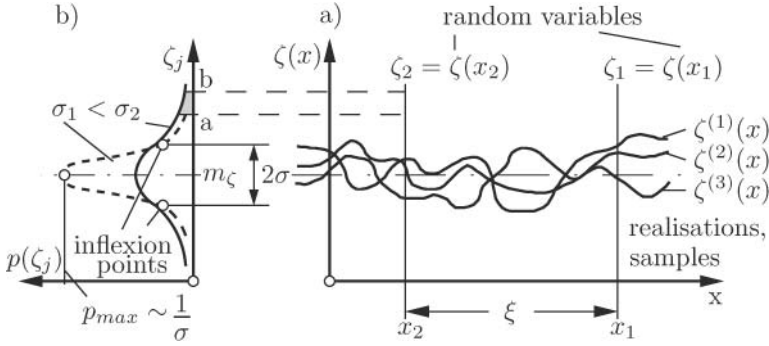
Let us consider a special type of road, e.g. the network of highways in a country, and let us assume that the roughness has been recorded by a large number of sectional profile measurements  $\zeta^{(r)}(x)$ ,  $r = 1, 2, \dots$ , that are put together in a diagram one above the others, see Figure 12a). The independent variable  $x$  denotes the distance from an arbitrary chosen starting point in each section. The single roughness profiles, also called realisations or samples, are different from each other,  $\zeta^{(r)}(x) \neq \zeta^{(s)}(x)$  for  $r \neq s$ . The complete set of measurement records, however, builds a random process  $\zeta(x)$  that represents the highway roughness. The ensemble of roughness values  $\zeta_j = \zeta(x_j)$  for a definite distance  $x = x_j$ ,  $j = 1, 2, \dots$ , result in a random variable. It can statistically be evaluated resulting in the probability density function  $p(\zeta)$ , see Figure 12b). From  $p(\zeta)$  follows the probability  $\text{Pr}$  that the random variable  $\zeta_j$  stays within a certain interval  $[a, b]$ ,

$$\text{Pr}(a \leq \zeta_j \leq b) = \int_a^b p(\zeta_j) d\zeta_j, \quad 0 \leq \text{Pr} \leq 1. \quad (5.1)$$

If the interval grows to infinite then the roughness value  $\zeta_j$  is for sure in it, i.e., the probability is  $\text{Pr} = 1$ ,

$$\int_{-\infty}^{\infty} p(\zeta_j) d\zeta_j = 1. \quad (5.2)$$

The moments of the probability density function are called mathematical expectation  $E\{\cdot\}$ , they characterize the random variable. The most important ones are the first order



**Figure 12.** Guideway roughness represented by a stochastic process:  
a) realisations  $\zeta^{(r)}(x)$  and random variables  $\zeta_j$     b) probability density  $p(\zeta_j)$

moment called mean  $m_\zeta(x_j)$  or expected value, and the second order moment called mean square value  $m_{\zeta^2}(x_j)$ . For  $j = 1$  it follows

$$m_\zeta(x_1) = E\{\zeta(x_1)\} \equiv \int_{-\infty}^{\infty} \zeta_1 p(\zeta_1) d\zeta_1, \quad (5.3)$$

$$m_{\zeta^2}(x_1) = E\{\zeta^2(x_1)\} \equiv \int_{-\infty}^{\infty} \zeta_1^2 p(\zeta_1) d\zeta_1. \quad (5.4)$$

An important statistical parameter is the variance  $P_\zeta(x_1) \equiv \sigma_\zeta^2(x_1)$  or the square of the standard deviation  $\sigma_\zeta(x_1)$ . It is the ensemble average of the square of the deviation from the mean, a so called central second moment,

$$P_\zeta(x_1) \equiv \sigma_\zeta^2(x_1) = E\{[\zeta(x_1) - m_\zeta(x_1)]^2\} = E\{\zeta^2(x_1)\} - m_\zeta^2(x_1). \quad (5.5)$$

The expectation  $E\{\cdot\}$  is mathematically a linear operation, where  $E\{c\} = c$  for  $c = \text{const.}$  This property has been used in (5.5). First and second order moments are well known from the determination of the center of mass and the moment of inertia.

For the reader not familiar with the concept of mean and variance the mechanical analogy may be helpful. A rod or wire with infinite length that has a variable mass density  $\rho(x)$  (mass per unit length) identical to the probability density  $p(x)$  in the probability distribution. According to (5.2) the total mass of the rod is normalized to unity. The mean value (5.3) corresponds to the coordinate of the center of mass and the mean square value (5.4) can be interpreted as moment of inertia related to the origin. The variances (5.5), however, corresponds to the moment of inertia related to the center of mass. Thus, the relation  $\sigma_\zeta^2 = E\{\zeta^2\} - m_\zeta^2$  in (5.5) can be interpreted as the parallel-axes theorem (2.2). The first two moments contain essential information about the distribution of the corresponding quantities, i.e. probability and mass, respectively.



Now two random variables  $\zeta_1 = \zeta(x_1)$  and  $\zeta_2 = \zeta(x_2)$  of the random process  $\zeta$  are considered. From the statistical relationship between these variables insight is gained into the random process itself. This is achieved by the (auto) correlation  $R_\zeta(x_1, x_2)$  or the (auto) covariance  $P_\zeta(x_1, x_2)$ , where in the latter the central second moments are used,

$$R_\zeta(x_1, x_2) = E\{[\zeta(x_1)\zeta(x_2)]\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \zeta_1 \zeta_2 p(\zeta_1, \zeta_2) d\zeta_1 d\zeta_2, \quad (5.6)$$

$$\begin{aligned} P_\zeta(x_1, x_2) &= E\{[\zeta(x_1) - m_\zeta(x_1)][\zeta(x_2) - m_\zeta(x_2)]\} \\ &= E\{\zeta(x_1)\zeta(x_2)\} - m_\zeta(x_1)m_\zeta(x_2). \end{aligned} \quad (5.7)$$

Here,  $p(\zeta_1, \zeta_2)$  denotes the second order probability density function. The prefix "auto" refers to the fact that both random variables belong to the same random process. Otherwise, i.e. if the two random variables belong to different random processes, the prefix "cross" will be used. In the limit case  $x_1 = x_2$  of equal positions the correlation (5.6) becomes equal to the mean square value (5.4) and the covariance (5.7) is identical to the variance (5.5).

After these prerequisites the properties stationarity, normal or Gaussian distribution and ergodicity shall be explained and described mathematically.

**Stationarity.** A random process is called stationary (or homogeneous) if its statistical properties are invariant under a shift of the origin of the  $x$ -axis. Thus, all first order probability density functions are equal,  $p(\zeta_1) = p(\zeta_2) = p(\zeta)$ , and the second order probability density function  $p(\zeta_1, \zeta_2)$  only depends on the distance  $\xi = x_1 - x_2$ , known as correlation width. As a consequence, all random variables and, thus, the whole process have the same mean and the same mean square value, respectively. A stationary random process can always be centralized, so that its mean value becomes zero. Thus, for stationary processes it follows from (5.3) to (5.7)

$$m_\zeta(x) = \text{const} = 0, \quad (5.8)$$

$$R_\zeta(x_1, x_2) = P_\zeta(x_1, x_2) = R_\zeta(\xi = x_1 - x_2), \quad (5.9)$$

$$m_{\zeta^2}(x) = \sigma_\zeta^2(x) = R_\zeta(\xi = 0) = \text{const}. \quad (5.10)$$

For centered processes correlation and covariance are identical. According to (5.9) the correlation depends on the correlation width  $\xi$  only, thus, it is called correlation function and characterizes the entire random process.

**Normal distribution.** A random process is called normal distributed or a Gaussian process if for every value  $x_j$  the corresponding random variable  $\zeta_j = \zeta(x_j)$  has a probability density function

$$p(\zeta_j) = \frac{1}{\sigma_j \sqrt{2\pi}} \exp \left[ -\frac{(\zeta_j - m_j)^2}{2\sigma_j^2} \right], \quad (5.11)$$

where  $m_j = m_\zeta(x_j)$  and  $\sigma_j = \sigma_\zeta(x_j)$ . A normal distributed process can uniquely be defined by its mean and variance, i.e. by its first two moments. The higher order moments can be expressed by these two quantities. Stationary normal distributed processes are also independent from distance  $x_j$ . The graph of the probability density function (5.11) has the well known bell shape, see Figure 12. This curve shows a maximum value of  $p_{\max} = 1/(\sigma_j\sqrt{2\pi})$  at  $\zeta_j = m_j$  and is symmetrical with respect to  $\zeta_j = m_j$  with points of inflection at  $\zeta_j = m_j \pm \sigma_j = \pm\zeta^*$ , where  $p(\pm\zeta^*) = p_{\max}/\sqrt{e}$  holds. Small values of the standard deviation  $\sigma_j$  mean that the quantity  $\zeta_j$  is concentrated around  $m_j$ . From (5.1) the probability can be calculated, so that the normal distributed random variable  $\zeta_j$  stays in the interval of length  $2k\sigma_j$ ,  $k = 1, 2, \dots$ , around the mean  $m_j$ . These intervals are called confidential intervals or  $k\sigma$ -limits,  $k = 1, 2, \dots$ :

$$\begin{aligned}
 k = 1 : \quad \Pr(m - \sigma \leq \zeta \leq m + \sigma) &= \Pr(|\zeta - m| \leq \sigma) = 0,6827 , \\
 k = 2 : \quad &\Pr(|\zeta - m| \leq 2\sigma) = 0,9545 , \\
 k = 3 : \quad &\Pr(|\zeta - m| \leq 3\sigma) = 0,9973 , \\
 k = 4 : \quad &\Pr(|\zeta - m| \leq 4\sigma) = 0,99994 . \quad (5.12)
 \end{aligned}$$

Here, the index  $j$  has been dropped. The probability that a Gaussian random variable takes a value outside, e.g. the  $2\sigma$ -limit, is only  $\Pr = 0,0455 \hat{=} 4,55\%$ .

**Ergodicity.** A stationary random process is called ergodic if the ensemble averages of the random variables  $\zeta_j = \zeta(x_j)$  are equal to the corresponding spacial (or temporal) averages taken along the coordinate  $x$  (or  $t$ ) of any sufficiently long representative sample function  $\zeta^{(r)}(x)$ . This means in short

- ensemble average equals sample average.

For ergodic processes it follows:

$$m_\zeta = \lim_{X \rightarrow \infty} \frac{1}{2X} \int_{-X}^X \zeta^{(r)}(x) dx , \quad (5.13)$$

$$R_\zeta(\xi) = \lim_{X \rightarrow \infty} \frac{1}{2X} \int_{-X}^X \zeta^{(r)}(x)\zeta^{(r)}(x - \xi) dx , \quad (5.14)$$

$$m_{\zeta^2} = R_\zeta(\xi = 0) = \lim_{X \rightarrow \infty} \frac{1}{2X} \int_{-X}^X [\zeta^{(r)}(x)]^2 dx , \quad (5.15)$$

$$P_\zeta(\xi) = \lim_{X \rightarrow \infty} \frac{1}{2X} \int_{-X}^X [\zeta^{(r)}(x) - m_\zeta][\zeta^{(r)}(x - \xi) - m_\zeta] dx , \quad (5.16)$$

$$\sigma_\zeta^2 = P_\zeta(\xi = 0) = \lim_{X \rightarrow \infty} \frac{1}{2X} \int_{-X}^X [\zeta^{(r)}(x) - m_\zeta]^2 dx . \quad (5.17)$$

Often it is difficult to decide if a stationary random process is ergodic or not. In the present case of guideway irregularities ergodicity is at least plausible. With respect to the information that can be gained it is equivalent to analyze a single very long measurement record (sample function) of e.g. highway surface irregularities or one takes the ensemble of pieces of this record and performs a statistical analysis. In general the averaging of sample functions is less time consuming than the ensemble averaging. However, in practical applications neither an infinitely large ensemble nor a sample function of infinite length is given. So, both procedures results in approximations for the averages only.

In the following always centered random processes with mean  $m_\zeta = 0$  are assumed. Thus, there is no difference between mean square and variance or between correlation function and covariance function.

In engineering applications stationary processes are often characterized by their power spectral density (PSD) or shortly spectral density  $S_\zeta(\Omega)$ . It can be gained from the correlation function  $R_\zeta(\xi)$  by a Fourier transform,

$$S_\zeta(\Omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} R_\zeta(\xi) e^{-i\Omega\xi} d\xi, \quad (5.18)$$

$$R_\zeta(\xi) = \int_{-\infty}^{\infty} S_\zeta(\Omega) e^{i\Omega\xi} d\Omega, \quad (5.19)$$

$$m_\zeta = 0 : P_\zeta = \sigma_\zeta^2 = R_\zeta(\xi = 0) = \int_{-\infty}^{\infty} S_\zeta(\Omega) d\Omega. \quad (5.20)$$

Here,  $\Omega$  with unit rad/m denotes a spacial circular frequency (wave number) that is connected to the wave length  $\lambda$  and the spacial frequency  $F$  by  $\Omega = 2\pi / \lambda = 2\pi F$ . Eqs. (5.18) and (5.19) are also known as Wiener-Chintschin relation, however, in literature the factor  $1/2\pi$  can also be found in front of the integral in (5.19) or as factors  $1/\sqrt{2\pi}$  in both equations. The benefit of the notation chosen here is clear from (5.20), where the variance is equal to the spectral power. The random process  $\zeta$  of guideway irregularities can equivalently be characterized by the correlation function  $R_\zeta(\xi)$  in spacial domain or by the spectral density  $S_\zeta(\Omega)$  in frequency domain. The variance  $\sigma_\zeta^2$  follows from  $R_\zeta(\xi)$  for  $\xi = 0$  or as integral over the variance fractions  $S_\zeta(\Omega) d\Omega$  related to the frequency  $\Omega$ . Such a fraction can also be interpreted as variance after the process has passed a narrow band filter of bandwidth  $d\Omega$ . Both functions  $R_\zeta(\xi)$  and  $S_\zeta(\Omega)$  are even functions,

$$R_\zeta(\xi) = R_\zeta(-\xi), \quad S_\zeta(\Omega) = S_\zeta(-\Omega). \quad (5.21)$$

This property allows an easy computation of the integrals in eqs. (5.14), (5.18) - (5.20) and leads to the introduction of the single sided spectral density  $\Phi_\zeta(\Omega)$  with nonnegative circular frequencies

$$\Phi_\zeta(\Omega) = \begin{cases} 2S(\Omega) & \Omega \geq 0 \\ 0 & \text{for } \Omega < 0 \end{cases}. \quad (5.22)$$

Correspondingly,  $S(\Omega)$  is sometimes denoted as double sided spectral density. Table 3 shows some often used correlation functions  $R_\zeta(\xi)$  and their corresponding single sided spectral densities  $\Phi_\zeta(\Omega)$ .

**White and coloured noise.** The correlation functions V, I and IV of Table 3 are used to derive and interpret the corresponding power spectral densities (PSD).

First, the Fourier transform for the single sided PSD  $\Phi(\Omega)$  will be derived. Introducing Euler's formula  $\exp(\pm i\Omega\xi) = \cos \Omega\xi \pm i \sin \Omega\xi$  into (5.18), (5.19) and splitting the range of integration yields for (5.18)

$$S(\Omega) = \frac{1}{2\pi} \left[ \int_{-\infty}^0 R(\xi)(\cos \Omega\xi - i \sin \Omega\xi) d\xi + \int_0^{\infty} R(\xi)(\cos \Omega\xi - i \sin \Omega\xi) d\xi \right]. \quad (5.23)$$

Using  $\xi = -\bar{\xi}$  in the first integral, where  $\bar{\xi}$  is a new variable, and regarding  $R(\xi) = R(-\xi) = R(\bar{\xi})$  according to (5.21) results in

$$2S(\Omega) = \frac{1}{\pi} \left[ \int_0^{\infty} R(\bar{\xi})(\cos \Omega\bar{\xi} + i \sin \Omega\bar{\xi}) d\bar{\xi} + \int_0^{\infty} R(\xi)(\cos \Omega\xi - i \sin \Omega\xi) d\xi \right]. \quad (5.24)$$

Putting the right hand side together and regarding  $2S(\Omega) = \Phi(\Omega)$  it follows

$$\Phi(\Omega) = \frac{2}{\pi} \int_0^{\infty} R(\xi) \cos \Omega\xi d\xi. \quad (5.25)$$

The inverse transform follows in the same manner,

$$R(\xi) = \int_0^{\infty} \Phi(\Omega) \cos \Omega\xi d\Omega. \quad (5.26)$$

Introducing the three correlation functions from Table 3 into (5.25) yield the following results that are visualized in Figure 13.

1. Correlation function V,  $R(\xi) = q_w \delta(\xi)$  : The evaluation of (5.25) yields

$$\Phi(\Omega) = \frac{2}{\pi} \int_0^{\infty} q_w \delta(\xi) \cos \Omega\xi d\xi = \frac{q_w}{\pi} = \text{const}. \quad (5.27)$$

For the computation of the integral in (5.27) it has to be considered that the lower limit of the integral coincides with the coordinate where the Dirac distribution  $\delta(\xi)$  is nonzero, thus, not the entire integrand but only one half of it is filtered out by the integration. The same results follows, if the transform (5.18) is applied, where  $S(\Omega) = \Phi(\Omega)/2$  has to be introduced. As a consequence, the random process V that is completely uncorrelated and has a constant PSD for all frequencies. Such

**Table 3.** Correlation function  $R(\xi)$  and spectral densities  $\Phi(\Omega)$ 

No	Correlation Function $R(\xi) = \int_0^{\infty} \Phi(\Omega) \cos \Omega \xi d\Omega$	Single Sided Spectral Density $\Phi(\Omega) = \frac{2}{\pi} \int_0^{\infty} R(\xi) \cos \Omega \xi d\xi$
I	$\sigma^2 e^{-\alpha \xi }$	$\frac{2\alpha\sigma^2}{\pi} \cdot \frac{1}{\Omega^2 + \alpha^2}$
II	$\sigma^2 e^{-\alpha \xi } \cos \beta\xi$	$\frac{2\alpha\sigma^2}{\pi} \cdot \frac{\Omega^2 + \alpha^2 + \beta^2}{(\Omega^2 - \alpha^2 - \beta^2) + \alpha^2\Omega^2}$
III	$\sigma^2 e^{-\alpha \xi } (\cos \beta\xi - \frac{\alpha}{\beta} \sin \beta \xi )$	$\frac{4\alpha\sigma^2}{\pi} \cdot \frac{\Omega^2}{(\Omega^2 - \alpha^2 - \beta^2) + \alpha^2\Omega^2}$
IV	$\sigma^2 e^{-\alpha \xi } (\cos \beta\xi + \frac{\alpha}{\beta} \sin \beta \xi )$	$\frac{4\alpha\sigma^2}{\pi} \cdot \frac{\alpha^2 + \beta^2}{(\Omega^2 - \alpha^2 - \beta^2) + \alpha^2\Omega^2}$
V	$q_w \delta(\xi)$	$\frac{1}{\pi} q_w$ $\Omega_d = \sqrt{1 - D^2} \cdot \Omega_0$
VI	$q_w \frac{\sin \alpha\xi}{\pi\xi}$	Band Limited White Noise $\frac{1}{\pi} q_w$ if $0 \leq \Omega \leq \alpha$ 0 else
VII	$2q_w \frac{\sin \alpha\xi/2 \cdot \cos \beta\xi}{\pi\xi}$	Ideal Band Pass $\frac{1}{\pi} q_w$ if $0 \leq \beta - \frac{\alpha}{2} \leq \Omega \leq \beta + \frac{\alpha}{2}$ 0 else

a process is called white noise  $w(x)$ , because it contains a constant spectrum like white light. However, this noise process has an infinite variance,

$$P_w = \sigma_w^2 = R_w(\xi = 0) = \int_0^{\infty} \Phi(\Omega) \, d\Omega = \infty. \quad (5.28)$$

This shows that white noise is a mathematical idealization and does not occur in nature. However, it is a useful limit case of real noise and it can mathematically be handled easily, thus, it is applied very often. Stationary Gaussian white noise with zero mean,  $m_w = 0$ , and intensity  $q_w$  will be denoted by

$$w(x) \sim N(0, q_w). \quad (5.29)$$

2. Correlation function I,  $R(\xi) = \sigma^2 e^{-\alpha|\xi|}$ ,  $\alpha > 0$  : The solution requires the evaluation of 5.25,

$$\begin{aligned} \Phi(\Omega) &= \frac{2\sigma^2}{\pi} \int_0^{\infty} e^{-\alpha\xi} \cos \Omega\xi \, d\xi = \frac{2\sigma^2}{\pi} \left[ \frac{e^{-\alpha\xi}}{\Omega^2 + \alpha^2} (-\alpha \cos \Omega\xi + \Omega \sin \Omega\xi) \right]_{\xi=0}^{\xi=\infty}, \\ \Phi(\Omega) &= \frac{2\sigma^2}{\pi} \frac{\alpha}{\Omega^2 + \alpha^2}. \quad (5.30) \end{aligned}$$

Here, the computation is very simple, since  $|\xi| = \xi$  for  $\xi \geq 0$  holds and the right hand side of (5.30) vanishes at the upper limit. The result is a coloured noise process with a maximum PSD of  $\Phi_{\max} = 2\sigma^2/(\pi\alpha)$  at the spacial frequency  $\Omega = 0$  and it follows an asymptotic behaviour  $\Omega \sim 1/\Omega^2$  for  $\Omega \rightarrow \infty$ . Here, the variance  $\sigma^2 = R(\xi = 0)$  is finite. The correlation decreases with increasing distance  $\xi$  monotonically. Figure 13 shows clearly the differences to a white noise process.

3. Correlation function IV,  $R(\xi) = \sigma^2 e^{-\alpha|\xi|} (\cos \beta\xi + \frac{\alpha}{\beta} \sin \beta|\xi|)$ ,  $\alpha > 0$  : In this case, the solution is analogous to 2) regarding trigonometrical transformations for products of harmonic functions,

$$\begin{aligned} \Phi(\Omega) &= \frac{2\sigma^2}{\pi} \int_0^{\infty} e^{-\alpha\xi} (\cos \beta\xi + \frac{\alpha}{\beta} \sin \beta\xi) \cos \Omega\xi \, d\xi \\ &= \frac{2\sigma^2}{\pi} \int_0^{\infty} e^{-\alpha\xi} \frac{1}{2} [\cos(\beta + \Omega)\xi + \cos(\beta - \Omega)\xi] \\ &\quad + e^{-\alpha\xi} \frac{\alpha}{2\beta} [\sin(\beta + \Omega)\xi + \sin(\beta - \Omega)\xi] \, d\xi \end{aligned}$$

$$\begin{aligned}
\Phi(\Omega) &= \frac{\sigma^2}{\pi} \left\{ e^{-\alpha\xi} \left[ \frac{1}{\alpha^2 + (\beta + \Omega)^2} (-\alpha \cos(\beta + \Omega)\xi + (\beta + \Omega) \sin(\beta + \Omega)\xi) \right. \right. \\
&\quad + \frac{1}{\alpha^2 + (\beta - \Omega)^2} (-\alpha \cos(\beta - \Omega)\xi + (\beta - \Omega) \sin(\beta - \Omega)\xi) \\
&\quad + \frac{\alpha}{\beta(\alpha^2 + (\beta + \Omega)^2)} (-\alpha \sin(\beta + \Omega)\xi - (\beta + \Omega) \cos(\beta + \Omega)\xi) \\
&\quad \left. \left. + \frac{\alpha}{\beta(\alpha^2 + (\beta - \Omega)^2)} (-\alpha \sin(\beta - \Omega)\xi - (\beta - \Omega) \cos(\beta - \Omega)\xi) \right] \right\}_{\xi=0}^{\xi=\infty} \\
&= \frac{\sigma^2}{\pi} \left[ \frac{\alpha}{\alpha^2 + (\beta + \Omega)^2} + \frac{\alpha}{\alpha^2 + (\beta - \Omega)^2} \right. \\
&\quad \left. + \frac{\alpha}{\beta} \left( \frac{\beta + \Omega}{\alpha^2 + (\beta + \Omega)^2} + \frac{\beta - \Omega}{\alpha^2 + (\beta - \Omega)^2} \right) \right] \\
\Phi(\Omega) &= \frac{4\alpha\sigma^2}{\pi} \frac{\alpha^2 + \beta^2}{N} . \tag{5.31}
\end{aligned}$$

In the last expression the denominator can be expressed in different manner,

$$\begin{aligned}
N &= [\alpha^2 + (\beta + \Omega)^2][\alpha^2 + (\beta - \Omega)^2] \\
&= (\Omega^2 - \alpha^2 - \beta^2)^2 + 4\alpha^2\Omega^2 = (\Omega^2 + \alpha^2 - \beta^2)^2 + 4\alpha^2\beta^2 . \tag{5.32}
\end{aligned}$$

Again the result is a coloured noise process with the following maximum PSD  $\Phi_{\max}$ ,

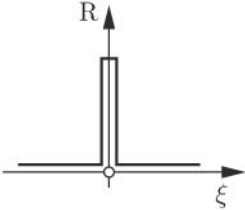
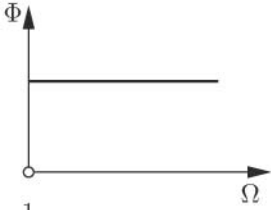
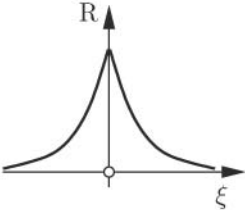
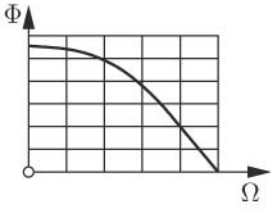
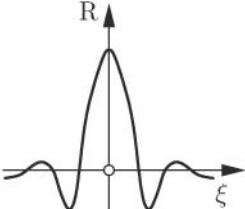
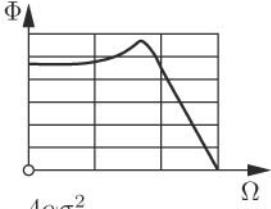
$$\Phi_{\max} = \begin{cases} \frac{\alpha\sigma^2}{\pi} \frac{\alpha^2 + \beta^2}{\alpha^2\beta^2} & \Omega = \sqrt{\beta^2 - \alpha^2} & \beta \geq \alpha \\ \frac{\alpha\sigma^2}{\pi} \frac{4}{\alpha^2 + \beta^2} & \Omega = 0 & \beta \leq \alpha \end{cases} \text{ at } \text{for} \tag{5.33}$$

It results in an asymptotic behaviour  $\Phi \sim 1/\Omega^4$  for  $\Omega \rightarrow \infty$ . The variance  $\sigma^2 = R(\xi = 0)$  is finite. Here, the correlation decreases with increasing distance  $\xi$  in an oscillating manner around zero, see Figure 13.

Until now only a scalar random process has been considered. If two scalar processes  $\zeta_{l,r}$  are given, for example the irregularities of the left and right track of a road, then the auto- and cross-correlation functions and their corresponding spectra have to be distinguished. This is usually being done by double indices of the cross correlation functions, where the Fourier transform is defined analogously to (5.18), (5.19),

$$S_{lr}(\Omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} R_{lr}(\xi) e^{-i\Omega\xi} d\xi , \tag{5.34}$$

$$R_{lr}(\xi) = \int_{-\infty}^{\infty} S_{lr}(\Omega) e^{i\Omega\xi} d\Omega . \tag{5.35}$$

name of noise process	correlation function $R(\xi)$	spectral density $\Phi(\Omega)$ (double log. plotted)
white noise	 $R = q_w \cdot \delta(\xi)$	 $\Phi = \frac{1}{\pi} \cdot q_w$
coloured noise I	 $R = \sigma^2 e^{-\alpha \xi }$	 $\Phi = \frac{2\alpha\sigma^2}{\pi} \cdot \frac{1}{\alpha^2 + \Omega^2}$
coloured noise IV	 $R = \sigma^2 e^{-\alpha \xi } \cdot (\cos \beta\xi + \frac{\alpha}{\beta} \sin \beta \xi )$	 $\Phi = \frac{4\alpha\sigma^2}{\pi} \cdot \frac{\alpha^2 + \beta^2}{(\alpha^2 - \beta^2 + \Omega^2)^2 + 4\alpha^2\beta^2}$

**Figure 13.** Correlation function  $R(\xi)$  and spectral density  $\Phi(\Omega)$  for white and coloured noise



The cross spectral densities are in general complex functions, where the following symmetries hold,

$$R_{lr}(\xi) = R_{rl}(-\xi) , \quad S_{lr}(\Omega) = S_{rl}(-\Omega) . \quad (5.36)$$

A normalized measure for the correlation between two processes is given by the coherence function  $\gamma(\Omega)$ ,

$$\gamma^2(\Omega) = \frac{|S_{lr}(\Omega)|^2}{S_l(\Omega)S_r(\Omega)} , \quad 0 \leq \gamma \leq 1 . \quad (5.37)$$

For  $\gamma \equiv 0$  both processes, e.g. the irregularities of left and right track, are completely uncorrelated, and for  $\gamma \equiv 1$  they are completely correlated. In general, the correlation is frequency dependent, as can be seen from (5.37).

If  $n$  scalar processes  $\zeta_v$ ,  $v = 1(1)n$ , are given, then it is suitable to combine them to a  $n \times 1$ -vector process  $\zeta$ . The mean  $m_\zeta$  is then replaced by the  $n \times 1$ -mean vector  $\mathbf{m}_\zeta$  and the variance  $P_\zeta$  by the  $n \times n$ -covariance matrix  $\mathbf{P}_\zeta$ ,

$$\begin{aligned} \mathbf{m}_\zeta(x) &= \mathbb{E}\{\zeta(x)\} , \\ \mathbf{P}_\zeta(x) &= \mathbb{E}\{[\zeta(x) - \mathbf{m}_\zeta(x)][\zeta(x) - \mathbf{m}_\zeta(x)]^T\} = \mathbf{P}_\zeta^T(x) . \end{aligned} \quad (5.38)$$

The variances of the single scalar processes appear as diagonal elements of the matrix  $\mathbf{P}_\zeta$ . It is clear that the variances are very important quantities to characterize the considered stationary, normal distributed and ergodic processes. The variances can be interpreted in different ways as

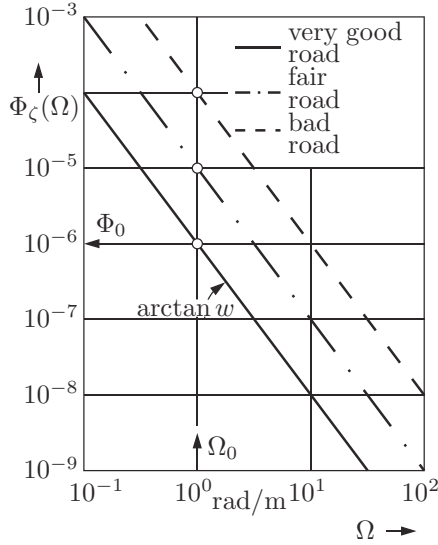
- central second order moment of a random variable, see (5.5), (5.38)
- mean square of a centered sample function, see (5.17)
- spectral power of a centered process, see (5.20)
- square of the standard deviation of the probability density function and, thus, defining the Gaussian bell shaped distribution (5.11) and the size of the confidence intervals (5.12).

## 5.2 Models of Road Roughness Profiles

The evaluation and preparation of numerous measurement data of road roughness profiles shows a tendency towards standardisation. A simple but often used road roughness model reads

$$\Phi_\zeta(\Omega) = \Phi_0 \left( \frac{\Omega_0}{\Omega} \right)^w , \quad 0 < \Omega_I \leq \Omega \leq \Omega_{II} < \infty , \quad (5.39)$$

where the frequency range is bounded. Here,  $\Omega_0$  [rad/m] denotes a standardised spacial circular frequency (wave number),  $\Phi_0 = \Phi_\zeta(\Omega_0)$  [m<sup>2</sup>/(rad/m)] characterizes the roughness at  $\Omega_0$  and can be considered as roughness level, and  $w$  is called waviness and characterizes whether the road contains long wave lengths ( $w$  large) or short wave lengths ( $w$  small) with pronounced spectral densities. For the wavelength  $\lambda$  the relation  $\lambda = 2\pi/\Omega$  holds. The waviness  $w$  of road profiles is found in the range  $1,75 \leq w \leq 2,25$  with  $w \approx 2$  as an average value. (5.39) is usually plotted in a  $\Omega, \Phi$ -diagram with double logarithmic scale and results in decreasing straight lines with slopes ( $-w$ ), see Figure 14.



**Figure 14.** Simple guideway roughness models for waviness  $w = 2$ ,  $\Phi_0 = \Phi(\Omega_0)$ ,  $\Omega_0 = 1$  rad/sec

A similar but more complicated model reads, Dodds and Robson (1973),

$$\Phi_\zeta(\Omega) = \begin{cases} \Phi_0 \left( \frac{\Omega_0}{\Omega} \right)^{w_1} & 0 < \Omega_I \leq \Omega \leq \Omega_0, \\ \Phi_0 \left( \frac{\Omega_0}{\Omega} \right)^{w_2} & \Omega_0 \leq \Omega \leq \Omega_{II} < \infty, \end{cases} \quad (5.40)$$

where different values  $w_1, w_2$ ,  $w_1 < w_2$ , of the waviness occur in the two frequency ranges.

The road roughness models (5.39) and (5.40) are approximations to measured PSD in a distinct frequency range  $0 < \Omega_I \leq \Omega \leq \Omega_{II} < \infty$ . In the limit case  $\Omega \rightarrow 0$  both models result in infinite values of the PSD,  $\Phi_\zeta(0) \rightarrow \infty$ , and, thus, in infinite variances,  $P_\zeta \rightarrow \infty$ . To avoid this unrealistic case, extended road roughness models have been suggested that hold in the entire frequency range. These models are coloured noise processes characterized by the PSD

$$\Phi_{\zeta_I}(\Omega) = \frac{2\alpha\sigma^2}{\pi} \frac{1}{\alpha^2 + \Omega^2}, \quad 0 \leq \Omega < \infty, \quad (5.41)$$

$$\Phi_{\zeta_{II}}(\Omega) = \frac{2\alpha\sigma^2}{\pi} \frac{\Omega^2 + \alpha^2 + \beta^2}{(\Omega^2 - \alpha^2 - \beta^2)^2 + 4\alpha^2\Omega^2}, \quad (5.42)$$

where  $\alpha$ ,  $\beta$  and  $\sigma^2$  are positive constants. The corresponding correlation functions can be calculated using

$$R_\zeta(\xi) = \int_0^\infty \Phi_\zeta(\Omega) \cos \Omega \xi \, d\Omega, \quad (5.43)$$

see Example 5.1. Inserting (5.41), (5.42) into (5.43) results in

$$R_{\zeta_I}(\xi) = \sigma^2 e^{-\alpha|\xi|}, \quad (5.44)$$

$$R_{\zeta_{II}}(\xi) = \sigma^2 e^{-\alpha|\xi|} \cos \beta \xi, \quad (5.45)$$

see also Figure 13, where  $\sigma^2 = R_\zeta(\xi = 0)$  denotes the finite variance of the road roughness profile. For  $\Omega \rightarrow \infty$  both PSDs (5.41), (5.42) decay according to  $\Phi \sim 1/\Omega^2$ . Thus, they approximate measured road roughness profiles very well.

### 5.3 Models of Vehicle Excitation

From the models of the road roughness profiles  $\zeta(x)$  shown in section 5.2 corresponding models of the vehicle excitation  $\zeta(t)$  in the time domain can be derived. Here, permanent road contact is assumed and, in a first step, only one contact point is considered. Transformation from space to time domain,  $\zeta(x) \rightarrow \zeta(t)$ , is performed considering the vehicle speed  $v(t)$ ,

$$dx(t) = v(t) dt, \quad x(t) = x(t_0) + \int_{t_0}^t v(\tau) d\tau. \quad (5.46)$$

From the road roughness profiles that are stationary in space domain it follows generally nonstationary vehicle excitation processes in time domain if the vehicle speed  $v(t)$  is time varying, Rill (1983) and Czerny (1987). This case will not be considered here. Throughout, we assume a constant vehicle speed,  $v = \text{const}$ , resulting in stationary vehicle excitation processes. For  $t_0 = 0$  and  $x(t_0) = 0$  from (5.46) it follows

$$x(t) = vt, \quad \xi = v\tau, \quad \omega = v\Omega, \quad v = \text{const}. \quad (5.47)$$

Here,  $\tau$  means the correlation time corresponding to the correlation width  $\xi$  and  $\omega$  with unit rad/s denotes the temporal circular frequency. Since the variances  $R_\zeta(0)$  of road profiles  $\zeta(x)$  and vehicle excitations  $\zeta(t)$  must be equal, (5.43) yields  $\Phi_\zeta(\omega) d\omega = \Phi_\zeta(\Omega) d\Omega$  that results with (5.47) in the PSD  $\Phi_\zeta(\omega)$  [ $\text{m}^2/(\text{rad/s})$ ] of the vehicle excitation

$$\Phi_\zeta(\omega) = \frac{1}{v} \Phi_\zeta \left( \Omega = \frac{\omega}{v} \right). \quad (5.48)$$

For example, from the road roughness model (5.39) with the average waviness  $w = 2$  it follows

$$\tilde{\Phi}_\zeta(\omega) = \frac{1}{v} \tilde{\Phi}_0 \left( \frac{v\Omega_0}{\omega} \right)^2 = v \tilde{\Phi}_0 \left( \frac{\Omega_0}{\omega} \right)^2. \quad (5.49)$$

Until now only stochastic excitation processes  $\zeta(t)$  of vehicle displacements were considered. But also the differentiated processes  $\dot{\zeta}(t)$  and  $\ddot{\zeta}(t)$  of vehicle velocity and vehicle acceleration, respectively, can be regarded. The corresponding PSDs read, Newland (1975),

$$\Phi_{\dot{\zeta}}(\omega) = \omega^2 \Phi_{\zeta}(\omega) , \quad \Phi_{\ddot{\zeta}}(\omega) = \omega^4 \Phi_{\zeta}(\omega) . \quad (5.50)$$

Eq. (5.50) applied to (5.49) for example yields a white noise process for the vehicle excitation velocity  $\dot{\zeta}(t)$  characterized by the PSD

$$\tilde{\Phi}_{\dot{\zeta}}(\omega) = \omega^2 v \tilde{\Phi}_0 \left( \frac{\Omega_0}{\omega} \right)^2 = v \tilde{\Phi}_0 \Omega_0^2 = \text{const} . \quad (5.51)$$

The corresponding correlation function reads

$$R_{\dot{\zeta}}(t) = q_{\dot{\zeta}} \delta(t) , \quad q_{\dot{\zeta}} = \pi v \tilde{\Phi}_0 \Omega_0^2 \quad (5.52)$$

where  $q_{\dot{\zeta}}$  denotes the noise intensity of  $\dot{\zeta}(t)$  and  $\delta(t)$  is the Dirac distribution. Assuming the waviness  $w = 4$  from (5.39) it follows similarly a white noise process for the vehicle excitation acceleration  $\ddot{\zeta}(t)$ ,

$$\tilde{\tilde{\Phi}}_{\ddot{\zeta}}(\omega) = \omega^4 v \tilde{\tilde{\Phi}}_0 \left( \frac{\Omega_0}{\omega} \right)^4 = v \tilde{\tilde{\Phi}}_0 \Omega_0^4 = \text{const} , \quad (5.53)$$

$$R_{\ddot{\zeta}}(t) = q_{\ddot{\zeta}} \delta(t) , \quad q_{\ddot{\zeta}} = \pi v \tilde{\tilde{\Phi}}_0 \Omega_0^4 . \quad (5.54)$$

The white noise processes are characterized by infinite variances and, thus, they are not realistic. On the other hand, assuming vehicle excitations like (5.51) to (5.54), the effort for the computation of vehicle random vibrations can be reduced considerably. Therefore, the white noise velocity excitation of road vehicles is important for engineering approximations.

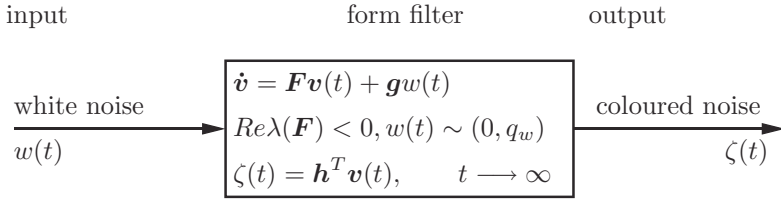
More realistic models of vehicle excitations  $\zeta(t)$  are based on (5.41) to (5.45), i. e., stationary Gaussian colored noise processes. These processes can be generated as steady state solution of a system of linear time invariant ordinary differential equations (ODEs) that are excited by white noise  $w(t)$ . These ODEs change - graphically speaking - the shape of the PSD or the corresponding correlation function, thus, they are called shape filter, see Figure 15.

Generally holds, if a stationary, ergodic Gaussian process excites a system of linear time invariant ODEs, then the properties stationarity, normal distribution and ergodicity will be transferred from the system input to the system output. The shape filter equations read

$$\dot{\zeta}(t) = \mathbf{h}^T \mathbf{v}(t) , \quad (5.55)$$

$$\dot{\mathbf{v}}(t) = \mathbf{F} \mathbf{v}(t) + \mathbf{g} w(t) , \quad \text{Re} \lambda(\mathbf{F}) < 0 , \quad w(t) \sim N(0, q_w) . \quad (5.56)$$

In steady state, the coloured noise process  $\zeta(t)$  results by superposition of the state variables  $v_i(t)$ ,  $i = 1(1)m$ , that are combined to the  $m \times 1$ -state vector  $\mathbf{v}$ . The shape filter has an asymptotically stable system matrix  $\mathbf{F}$  and the excitation is given by white



**Figure 15.** Blockdiagram of a shape filter for generation of coloured noise

noise  $w(t)$  with zero mean and intensity  $q_w$ . The quantities  $\mathbf{F}$ ,  $\mathbf{g}$  and  $\mathbf{h}$  characterize the shape filter completely. For the models (5.41) to (5.45) these quantities read

$$\text{I:} \quad \mathbf{F} = -\alpha v, \quad \mathbf{g} = g, \quad \mathbf{h} = 1, \quad (g^2 q_w = 2\alpha v \sigma^2), \quad (5.57)$$

$$\text{II:} \quad \mathbf{F} = \begin{bmatrix} 0 & 1 \\ -(\alpha^2 + \beta^2)v^2 & -2\alpha v \end{bmatrix}, \quad \mathbf{g} = g \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \\ \mathbf{h} = \begin{bmatrix} v\sqrt{\alpha^2 + \beta^2} \\ 1 \end{bmatrix}, \quad (g^2 q_w = 2\alpha v \sigma^2). \quad (5.58)$$

Since the product  $g^2 q_w$  of the shape filter quantities follows from a product of model parameters, there is a possibility of choice. If we put  $g = 1$  for example, then for both shape filters the intensity  $q_w$  of the white noise excitation reads  $q_w = 2\alpha v \sigma^2$ . This shows that the intensity of the vehicle excitation grows proportional to the vehicle speed  $v$ .

In general, the shape filter quantities should be determined directly from measured PSDs rather than from analytical approximations like shown here. However, in this case the order  $m$  of the filter has to be chosen. For  $m = 2$  this procedure is extensively described in Müller et al. (1980).

Until now only vehicle excitations caused by a single contact point have been considered resulting in a scalar stochastic excitation process. However, real multi-axle vehicles have multiple road contacts. For  $i = 1(1)q$  axles and  $k = 1(1)s$  tracks there are  $q \cdot s$  contact points resulting in correspondingly many scalar excitation processes that can be combined to a  $(q \cdot s) \times 1$ -vector process  $\boldsymbol{\zeta}$  of vehicle excitation. Then, the scalar relation (5.55) has to be replaced by the corresponding vector process

$$\boldsymbol{\zeta}(t) = \mathbf{H}\mathbf{v}(t) \quad (5.59)$$

where  $\mathbf{H}$  is a  $(q \cdot s) \times m$ -matrix, and the shape filter may excited by different white noise processes  $\mathbf{w}(t)$

$$\dot{\mathbf{v}} = \mathbf{F}\mathbf{v}(t) + \mathbf{G}\mathbf{w}(t), \quad \mathbf{w}(t) \sim \mathbf{N}(\mathbf{0}, \mathbf{Q}_w) \quad (5.60)$$

with the input intensity matrix  $\mathbf{Q}_w$

In case of subsequent contact points in a single track the excitation processes  $\zeta_i(t)$ ,  $i = 1(1)q$ , can be gained from the excitation process  $\zeta_1(t) \equiv \zeta(t)$  of the first contact point by

a time shift  $t_i$

$$\zeta_i(t) = \zeta(t - t_i), \quad t_i = \frac{l_i}{v}, \quad 0 = t_1 < t_2 < \dots < t_q, \quad i = 1(1)q \quad (5.61)$$

where  $l_i$  denotes the distance between the first and the  $i$ -th axle and  $v = \text{const}$  is the vehicle speed. Further details can be found in Müller and Popp (1979).

In case parallel tracks the statistical relationship between the parallel acting excitation processes have to be regarded. Measurements of correlation functions are reported in Bormann (1978). A corresponding shape filter is given by Rill (1983).

A typical vector process characterizing the excitation of a four wheel vehicle reads ( $f \hat{=}$  front,  $r \hat{=}$  rear,  $l \hat{=}$  left,  $r \hat{=}$  right),

$$\zeta(t, T) = \begin{bmatrix} \zeta_{lf}(t) \\ \zeta_{rf}(t) \\ \zeta_{lr}(t) \\ \zeta_{rr}(t) \end{bmatrix} = \begin{bmatrix} \zeta_l(t) \\ \zeta_r(t) \\ \zeta_l(t - T) \\ \zeta_r(t - T) \end{bmatrix}, \quad T \equiv t_2 = \frac{l_2}{v}, \quad (5.62)$$

where the delay time  $T$  follows from the axle distance  $l_2$  and the vehicle speed  $v$ .

## 6 Ride Comfort and Safety Criteria

The assessment of a vehicle's dynamical properties is related to its three essential motions:

- longitudinal motion (driving and braking),
- lateral motion (guidance and steering),
- vertical motion (suspension and damping).

These three motions of a vehicle are more or less decoupled, so that the assessment criteria can be formulated separately for each motion. Of course, coupled criteria can be developed if necessary. But then, one has to compete with the subjective impression during driving tests. Thus, the question of a mathematical formulation of assessment criteria is difficult and can often be answered empirically only. Sometimes an interdisciplinary approach is helpful taking into account ergonomical knowledge.

The criteria to assess the longitudinal motion are known as vehicle performance. It comprises maximum speed, resistance of motion and vehicle acceleration. These criteria are self explaining and do not need any further consideration. However, it should be mentioned that the exact measurement of the vehicle performance in driving test is a nontrivial problem.

The essential criterium for the lateral motion is the driving stability. The vehicle should safely follow a trajectory prescribed by the steering systems. Thus, a stability problem is given, that can be solved applying the well developed methods of system dynamics.

For the vertical motion the criteria ride comfort and safety are essential. The wheel suspension should absorb the disturbances coming from the road irregularities so that a comfortable ride becomes possible. At the same time the weight of the vehicle should be transferred to the road without wheel vibrations, so that large lateral forces in the contact area remain in order to guarantee safety.

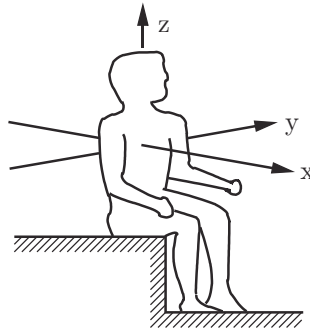
Besides the vehicle motions the static and dynamic forces acting on the vehicle parts are important criteria, because the life time of the vehicle components depends strongly on the loads. These loads can be calculated in advance by model based numerical simulations. In the following sections the criteria ride comfort and safety will be considered in detail.

### 6.1 Ride Comfort

Ride comfort is based on the subjective human perception. Numerous experimental investigations in ergonomics have shown that the human perception of vibrations depends on the acceleration,

$$K = K(a) , \quad (6.1)$$

where  $K$  is a nondimensional perception measure and  $a$  is the absolute value of the acceleration in horizontal or vertical direction, respectively, see Figure 16. Furthermore, the position of the human body (sitting, standing or lying) is of importance. In ergonomics the vibration tests have been performed using deterministic especially harmonic excitations. In vehicles, however, the human body usually is exposed to random vibrations. Therefore, in the following the different types of excitation will be investigated separately.



**Figure 16.** Direction of exposure to vibration of a seated human

**Deterministic excitation** For a rough qualitative assessment of ride comfort during a deterministic excitation

$$K \sim a_{\max} , \quad (6.2)$$

is assumed, i.e. the maximum acceleration value prevails the perception measure. As a thumb rule  $a_{\max} \leq 0.5 \text{ m/s}^2$  results in a good ride comfort. More refined relations between  $K$  and  $a$  are given in international or national standards that are based on extensive ergonomical investigations, see ISO International Standard 2631 (1974) or VDI Richtlinie 2057 (1975-1979).

For harmonic excitations the aforementioned standards give a precise relation that suits well for vehicle dynamics. Starting from the harmonic acceleration

$$a(t) = A \sin \omega t, \quad \omega = 2\pi f, \quad (6.3)$$

with frequency  $f$  [Hz] the root mean square (rms) value  $a_{\text{rms}}$  of the acceleration can be determined,

$$a_{\text{rms}} = \sqrt{\frac{1}{T} \int_0^T a^2(t) dt} = \frac{A}{\sqrt{2}} [\text{m/s}^2], \quad (6.4)$$

where the unit  $\text{m/s}^2$  has to be used. Then the perception measure  $K$  reads for vertical excitation

$$\begin{aligned} K &= 10a_{\text{rms}}\sqrt{f}, & 1 \leq f \leq 4, \\ K &= 20a_{\text{rms}}, & 4 \leq f \leq 8, \\ K &= 160a_{\text{rms}}f^{-1}, & 8 \leq f \leq 80, \end{aligned} \quad (6.5)$$

and for horizontal excitation

$$\begin{aligned} K &= 28a_{\text{rms}}, & 1 \leq f \leq 2, \\ K &= 56a_{\text{rms}}f^{-1}, & 2 \leq f \leq 80. \end{aligned} \quad (6.6)$$

Eqs. (6.5) and (6.6) can be interpreted as experimentally determined frequency response functions of the human perception, they are depicted in Figure 17.

From Figure 17a) can be seen that a vertical excitation between 4 and 8 Hz is perceived as very unpleasant, because in this range occurs resonance of the human stomach. The perception  $K$  is a nondimensional measure characterizing the subjective impression of test persons related to the criteria well-being and comfort, efficiency of labour and impairment of health, respectively. Here, the exposure time plays an essential role. The standards ISO 2631 and VDI 2057 describe the relation between exposure time and the mentioned criteria, see Figure 18. The perception measure required for road vehicles is about

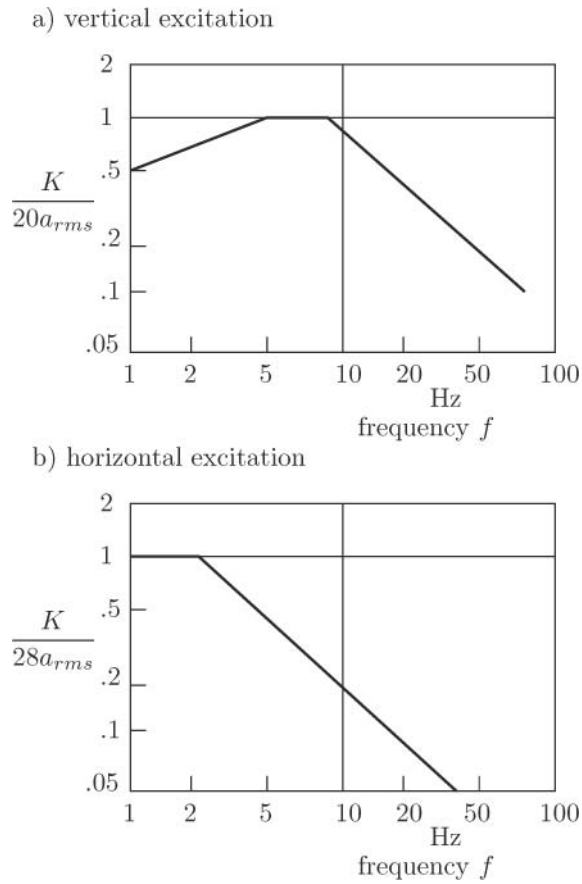
$$2 < K < 10. \quad (6.7)$$

Subclasses of the perception can be defined and the corresponding vibration levels can be described:

$$\begin{aligned} C1 / C2 &\text{ noticeable,} \\ D1 / D2 &\text{ strong noticeable,} \\ E1 / E4 &\text{ very strong noticeable.} \end{aligned} \quad (6.8)$$

Obviously, the description of perception shows some uncertainties, due to the subjective human nature.





**Figure 17.** Frequency responses of the perception measure:  
a) vertical exponse    b) horizontal exponse

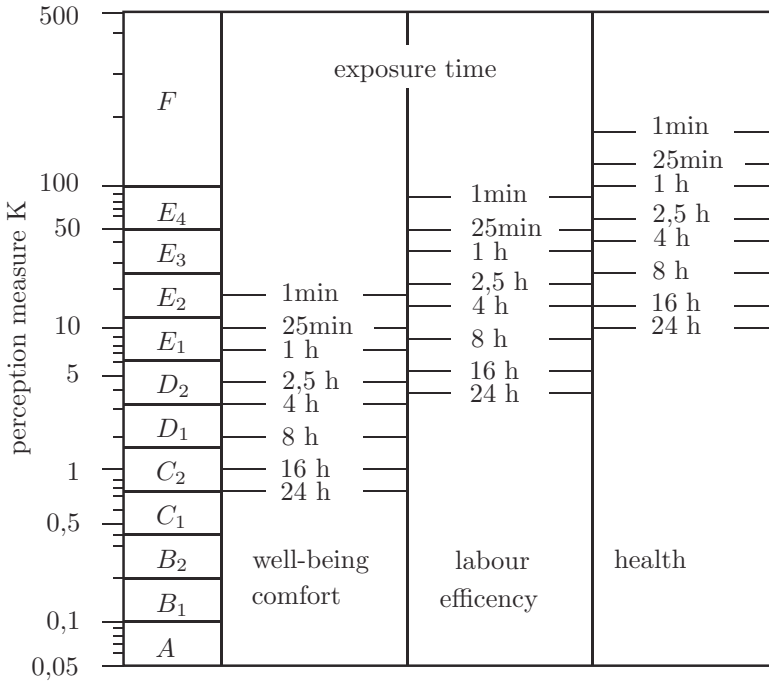


Figure 18. Perception measure and exposure time of vibrations

**Stochastic excitation** Experimental results for a stochastic excitation of test persons are very limited. However, they confirm the interpretation of the subjective human perception as response of a vibration system for a stochastic excitation, too, and the possibility to describe the perception using methods of linear system theory. The rms-value of a Gaussian process  $a(t)$  is equivalent to its standard deviation  $\sigma_a$ ,

$$a_{\text{rms}} = \sigma_a = \sqrt{\int_0^\infty \Phi_a(\omega) d\omega}, \tag{6.9}$$

where  $\Phi_a(\omega)$  is the single sided power spectral density (PSD) of the process under consideration. The standard deviation is a scalar parameter characterizing globally the stochastic process, however, it is frequency independent. Thus, the weighting of frequencies and the calculation of the rms-value have to be interchanged, which is allowed for linear stochastic systems. This results in the variance  $\sigma_{\bar{a}}^2$  of the weighted process  $\bar{a}(t)$ ,

$$\sigma_{\bar{a}}^2 = \int_0^\infty \alpha^2 |F(\omega)|^2 \Phi_a(\omega) d\omega. \tag{6.10}$$

Here,  $\alpha$  is a dimensional factor,  $F(\omega)$  is the frequency response function of a weighting filter and  $\Phi_a(\omega)$  is the single sided PSD of the mechanical acceleration process  $a(t)$ . The still unknown quantities of the frequency weighting in (6.10) can be found by comparison with the results for deterministic excitation. For vertical excitation the dimensional factor  $\alpha$  reads,

$$\alpha = 20 \text{ s}^2/\text{m} \quad (6.11)$$

and the frequency response function is given by

$$|F(\omega)| = \frac{1}{20} K(a_{\text{rms}}, f), \quad f = \omega/2\pi, \quad (6.12)$$

where the perception  $K$  according to (6.5) is used.

The frequency response function (6.12) can be well approximated by a linear weighting filter or shape filter. This yields in the frequency domain

$$\begin{aligned} F(\omega) &= \bar{\mathbf{h}}^T (i\omega \mathbf{E} - \bar{\mathbf{F}})^{-1} \bar{\mathbf{g}} \\ &= \frac{b_0 + b_1(i\omega) + \dots + b_r(i\omega)^r}{a_0 + a_1(i\omega) + \dots + a_{s-1}(i\omega)^{s-1} + (i\omega)^s}. \end{aligned} \quad (6.13)$$

Here,  $\bar{\mathbf{h}}$  and  $\bar{\mathbf{g}}$  are  $s \times 1$ -vectors and  $\bar{\mathbf{F}}$  is a  $s \times s$ -matrix characterizing the shape filter. Furthermore,  $a_i, i = 0(1)s - 1$ , and  $b_j, j = 0(1)r \leq s$  are coefficient of the shape-filter response function. The following relation holds,

$$\bar{\mathbf{F}} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{s-1} \end{bmatrix}, \quad \bar{\mathbf{g}} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}, \quad \bar{\mathbf{h}}^T = [b_0 \ b_1 \ \dots \ b_r \ \dots \ 0], \quad (6.14)$$

i.e. the system matrix  $\bar{\mathbf{F}}$  can be chosen as Frobenius matrix. In time domain the shape filter can be described equivalently by a differential equation,

$$\begin{aligned} \bar{a}(t) &= \alpha \bar{\mathbf{h}}^T \bar{\mathbf{v}}(t), \\ \dot{\bar{\mathbf{v}}}(t) &= \bar{\mathbf{F}} \bar{\mathbf{v}}(t) + \mathbf{g} a(t). \end{aligned} \quad (6.15)$$

Here,  $\bar{\mathbf{v}}$  is the  $s \times 1$ -state vector of the shape filter that is excited by the acceleration  $a(t)$ . The scalar product of the  $s \times 1$ -vectors  $\bar{\mathbf{h}}$  and  $\bar{\mathbf{v}}(t)$  results in the frequency weighted scalar acceleration  $\bar{a}(t)$ . The coefficients of the shape filter have to be determined in such a way that the given frequency response function (6.12) is approximated sufficiently well. This can always be achieved by choosing the order  $s$  of the filter large enough.

**Shape filter for the human perception.** A second order shape filter shall be determined so that the frequency response function (6.5) for the human perception of vibration is approximated well.

The shape filter of second order,  $s = 2$ , has the following form according to (6.14) and (6.15):

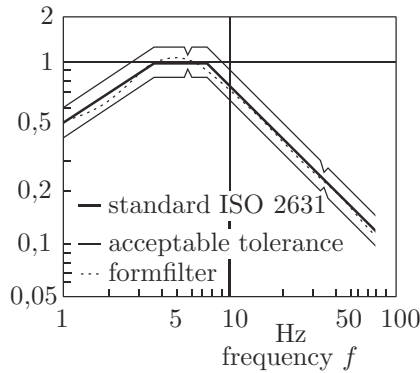
$$\bar{a} = \alpha [b_0 b_1] \begin{bmatrix} \bar{v}_1 \\ \bar{v}_2 \end{bmatrix}, \quad (6.16)$$

$$\begin{bmatrix} \dot{\bar{v}}_1 \\ \dot{\bar{v}}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -a_0 & -a_1 \end{bmatrix} \begin{bmatrix} \bar{v}_1 \\ \bar{v}_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} a(t). \quad (6.17)$$

Choosing the numbers

$$\begin{aligned} a_0 &= 1200s^{-2}, & b_0 &= 500s^{-2}, \\ a_1 &= 50s^{-1}, & b_1 &= 50s^{-1}, \\ \alpha &= 20s^2m^{-1} \end{aligned}$$

in case of vertical excitation results in the frequency response function depicted in Figure 19. It can be seen, that already a second order shape filter approximates the standard ISO 2631 surprisingly well. The deviations remain completely within the tolerances allowed by the standard.



**Figure 19.** Shape filter frequency response of perception in vertical direction

Similarly, a second order shape filter regarding the frequency weighting (6.6) for the horizontal excitation can be determined using the numbers

$$\begin{aligned} a_0 &= 75s^{-2}, & b_0 &= 31,25s^{-2}, \\ a_1 &= 12,5s^{-1}, & b_1 &= 12,5s^{-1}, \\ \alpha &= 28s^2m^{-1} \end{aligned}$$

Thus, the computation of the perception quantity  $K$  results simply in

$$K = \sigma_{\bar{a}}, \quad (6.18)$$

that includes the frequency weighting automatically. Using again the  $K$ -value from (6.18), the allowed exposure time for the criteria comfort, efficiency and health follows from Figure 18.

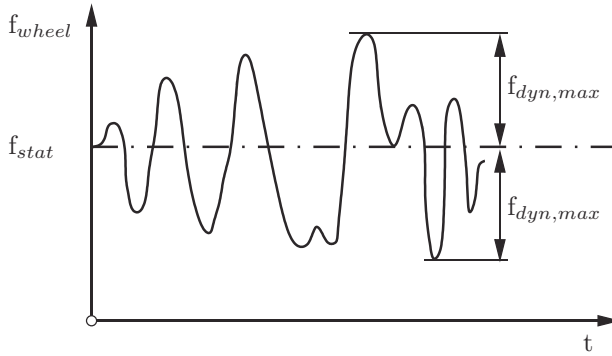
## 6.2 Ride Safety

Ride safety is related to the longitudinal braking forces and the lateral guidance forces of a vehicle. The tangential contact forces between tyre and road are influenced essentially by the normal forces in the contact area. Hence, the remaining minimum wheel load is a criterion for ride safety.

The total wheel load  $f_{\text{Wheel}}$  can be divided into a static load  $f_{\text{Stat}}$  due to the vehicle weight and a dynamic load  $f_{\text{Dyn}}$  generated by vehicle vibrations,

$$f_{\text{Wheel}}(t) = f_{\text{Stat}} + f_{\text{Dyn}}(t) , \quad (6.19)$$

visualized in Figure 20.



**Figure 20.** Statical and dynamical wheel load

The remaining minimum total wheel load reads

$$f_{\text{Wheel, min}} = f_{\text{Stat}} - f_{\text{Dyn, max}} . \quad (6.20)$$

This defines the safety margin  $R$ ,

$$R = \frac{f_{\text{Stat}} - f_{\text{Dyn,max}}}{f_{\text{Stat}}} , \quad (6.21)$$

a scalar quantity varying in the range of  $0 \leq R \leq 1$ . The safety margin for a vehicle in standstill is  $R = 1$ , it can drop for a fast moving vehicle on a rough road down to  $R = 0$ .

For roads with a randomly disturbed surface the dynamic wheel load variations can be determined from numerical simulations. For an assessment of the ride safety the standard deviation  $\sigma_f$  of the vertical wheel load variations can be used. This yields

$$R = \frac{f_{\text{Stat}} - \sigma_f}{f_{\text{Stat}}} , \quad (6.22)$$

where  $R$  has to be considered as a statistical quantity.

The computation of the dynamical wheel load variations, i.e.  $f_{Dyn}$  or  $\sigma_f$ , respectively, is based on the state and input quantities of the global vehicle guideway system. Depending on the tyre model also the relative displacements and velocities have to be known. Hence, the following general expressions can be given,

$$f_{Dyn} = \mathbf{e}^T \mathbf{x} \quad (6.23)$$

where  $\mathbf{e}$  denotes a  $n \times 1$ -vector of coefficients and  $\mathbf{x}(t)$  the  $n \times 1$ -state vector.

## 7 Computation of the Random Vehicle Response

The global vehicle system including the guideway excitation is composed of the vehicle model (3.11) or (3.12), respectively, and the guideway model (5.59) and (5.60). Considering only the excitation by random guideways  $\mathbf{u}_F(t) = \boldsymbol{\zeta}(t)$  it remains for linear motions

$$\begin{aligned} \begin{bmatrix} \dot{\mathbf{x}}_F \\ \dot{\mathbf{v}} \end{bmatrix} &= \underbrace{\begin{bmatrix} \mathbf{A}_F & \mathbf{B}_F \mathbf{H} \\ \mathbf{0} & \mathbf{F} \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{bmatrix} \mathbf{x}_F \\ \mathbf{v} \end{bmatrix}}_{\mathbf{x}} + \underbrace{\begin{bmatrix} \mathbf{0} \\ \mathbf{G} \end{bmatrix}}_{\mathbf{B}} \mathbf{w}, \\ \dot{\mathbf{x}} &= \mathbf{A} \mathbf{x} + \mathbf{B} \mathbf{w}. \end{aligned} \quad (7.1)$$

The entire vehicle system including the guideway excitation can be described accordingly by the nonlinear state equation

$$\dot{\mathbf{x}}(t) = \mathbf{a}(\mathbf{x}, \mathbf{w}, t), \quad (7.2)$$

where  $\mathbf{x}$  is the global state vector,  $\mathbf{w}$  characterizes the excitation, and  $\mathbf{a}$  denotes a nonlinear vector function. In the following we will assume that (7.2) can be linearized resulting in the linear state equation (7.1) again where in many cases the state matrix  $\mathbf{A}$  and the input matrix  $\mathbf{B}$  are time invariant.

The state equation serves as basis of the evaluation of the vehicle performance. However, in case of ride comfort evaluation another shape filter has to be added. The input of this shape filter is the acceleration  $a(t)$  at the seat position, that can be composed of the states and excitation quantities of the entire vehicle system,

$$a(t) = \mathbf{c}^T \mathbf{x}(t) + \mathbf{d}^T \mathbf{w}(t), \quad (7.3)$$

where  $\mathbf{c}$  is a weighting vector of the states and  $\mathbf{d}$  a weighting vector of the excitations. Thus, the state equations (7.1) extended by the shape filter (6.15) reads

$$\begin{aligned} \begin{bmatrix} \dot{\hat{\mathbf{x}}} \\ \dot{\hat{\mathbf{v}}} \end{bmatrix} &= \underbrace{\begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{g}\mathbf{c}^T & \mathbf{F} \end{bmatrix}}_{\tilde{\mathbf{A}}} \underbrace{\begin{bmatrix} \mathbf{x} \\ \mathbf{v} \end{bmatrix}}_{\tilde{\mathbf{x}}} + \underbrace{\begin{bmatrix} \mathbf{B} \\ \mathbf{g}\mathbf{d}^T \end{bmatrix}}_{\tilde{\mathbf{B}}} \mathbf{w}, \\ \dot{\hat{\mathbf{x}}} &= \tilde{\mathbf{A}} \tilde{\mathbf{x}} + \tilde{\mathbf{B}} \mathbf{w}. \end{aligned} \quad (7.4)$$

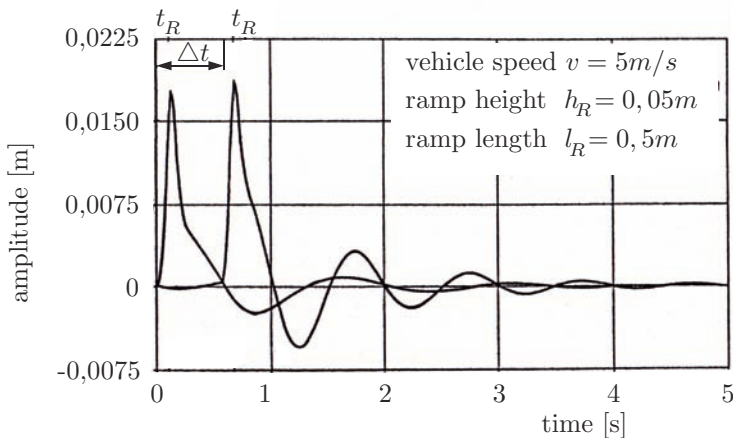
The extended system (7.4) has exactly the same structure as the vehicle system (7.1), so that we consider system (7.1) subsequently without loss of generality. The required computational methods can be subdivided into numerical and analytical methods that will be described in the following sections.

## 7.1 Numerical Simulation

The given set of linear or nonlinear differential equations (7.1), (7.2) can be analyzed by numerical integration. For random excitations a single realization is sufficient due to ergodicity. In case of nonlinear systems the random response in general is not normal distributed any more. Even so, in many cases one computes the first two moments only for evaluation purposes. However, the effort of numerical simulations is always large due to the complex vehicle system models. Thus, the proper choice of the integration methods is very important. Unfortunately, it is not possible to give general recommendations since on the one hand new integration methods will be developed by numerical mathematicians, on the other hand the performance of the computers still increases. Often the users applies the integration methods at hand and performs test runs for comparison.

**Simulation of vertical motions of vehicles.** Different integration methods are applied to the simulation of the vertical motion of the front and rear axle of a nonlinear 16 degree of freedom vehicle model under a ramp excitation. Furthermore, a vehicle with 5 degrees of freedom in vertical direction is considered as a test example of a random excitation.

The ramp is crossed first by the front axle and then by the rear axle, this results in a time delay in the corresponding vibration responses, see Figure 21.



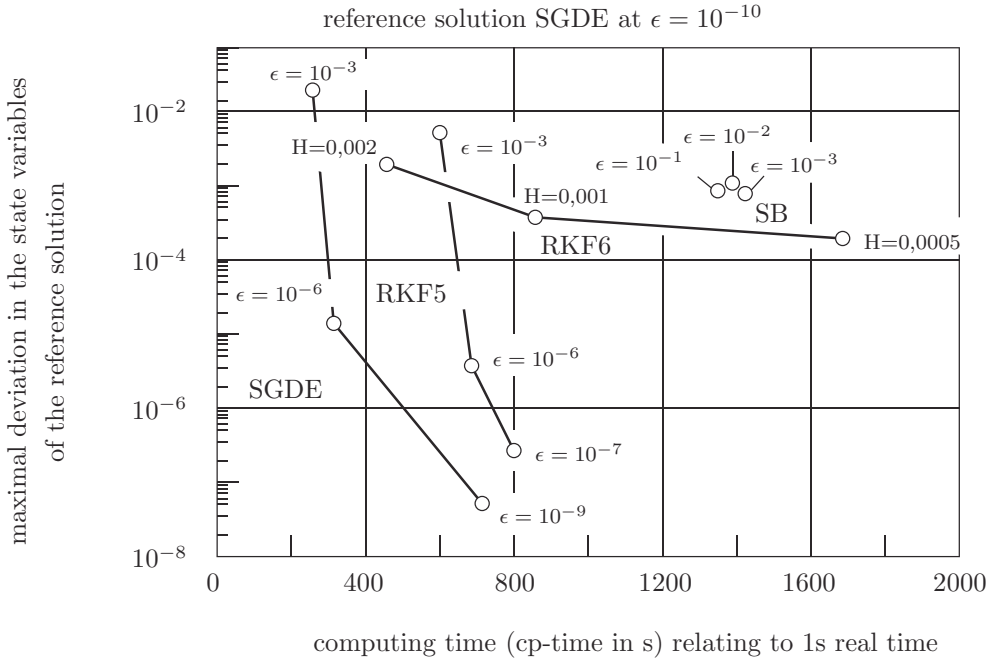
**Figure 21.** Vertical response of front and rear wheels of a 16-DOF vehicle model after ramp excitation

According to Rill (1981), the following time integration codes are compared:

- Single step methods (Runge-Kutta-Fehlberg procedure of order 5 and 6 - RKF5, RKF6 -),
- Predictor-corrector multi step method (Shampine-Gordon procedure - SGDE -),

- Extrapolation method (Stoer-Bulirsch procedure - SB -).

The results of the large vehicle model are shown in Figure 22. It can be seen that the Shampine-Gordon procedure (Shampine and Gordon, 1984) leads to the best results with respect to computing time and accuracy.

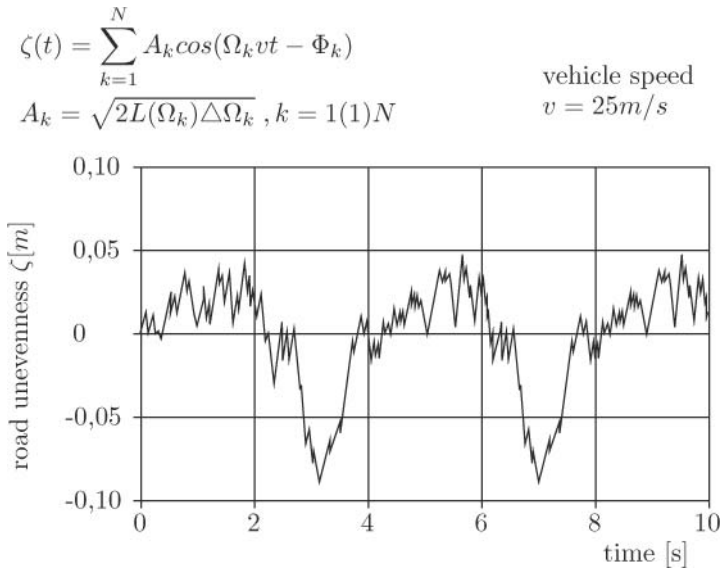


**Figure 22.** Comparison of time integration codes

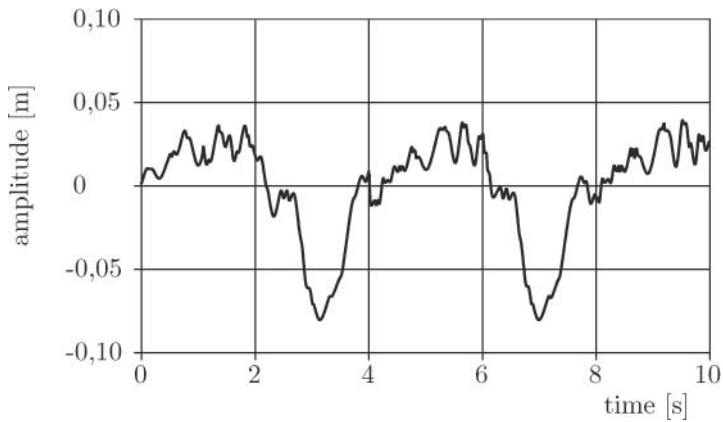
The random excitation of the small vehicle model is generated by a random superposition of harmonic functions characterizing the road roughness profile, see Figure 23. The vertical motion of the front axle, see Figure 24, follows the road profile, however, the high frequency components of the excitation are not visible. Nevertheless, the integration procedure must take into account the high frequencies. Applying the procedures mentioned above, lead to the results shown in Figure 25. Here, the simple Runge-Kutta-Fehlberg procedures are superior due to the fast evaluation of the right hand side of the differential equations.

The numerical integration procedures result in time histories that have to be evaluated in order to assess ride comfort and safety. In case of random excitation the computation of the performance criteria requires time averages that can be gained directly from the simulation results.





**Figure 23.** Realisation of a random process of road roughness



**Figure 24.** Vertical response of the front wheels of a nonlinear 5-DOF vehicle model during random excitation

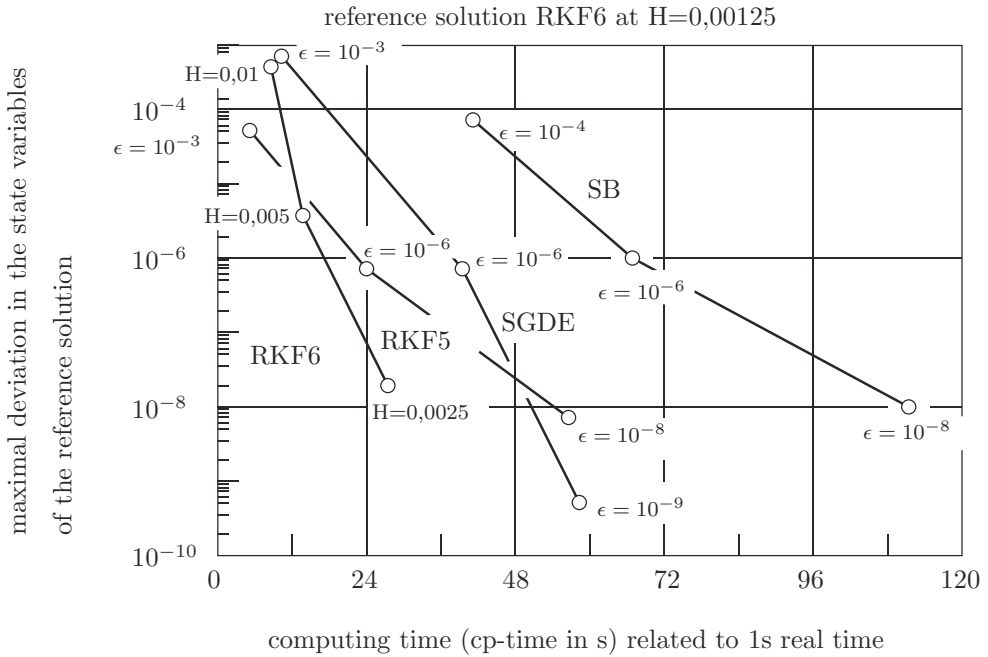


Figure 25. Comparison of time integration codes

## 7.2 Spectral Density Analysis

The investigation of random vibrations can be performed in frequency domain using spectral analysis or in time domain applying covariance analysis. Assuming that  $\mathbf{x}(t) = \mathbf{0}$  is an asymptotically stable equilibrium position, then from  $E \{ \mathbf{w}(t) \} = \mathbf{0}$  it follows  $E \{ \mathbf{x}(t) \} = \mathbf{0}$ , i. e. the mean value of the state vector vanishes in steady state. Thus, as essential goal it remains to calculate the characterizing variances, e. g. the variance  $\sigma_{\bar{a}}^2 = E \{ \bar{a}^2(t) \}$  of the frequency weighted acceleration  $\bar{a}(t)$ .

First we calculate the power spectral density (PSD) that is a characteristic quantity of a random process in frequency domain. In particular, for linear systems the PSD of the input and output processes are related to each other by its frequency response. However, one should keep in mind that the ride comfort and safety criteria depend primarily on variances and not on PSDs.

The  $n \times n$ -PSD-matrix  $\mathbf{S}_x(\omega)$  of the entire vehicle system described by the state equations (7.1), where  $\mathbf{A} = \mathbf{const}$  and  $\mathbf{B} = \mathbf{const}$  are assumed, follows from the white noise process  $\mathbf{w}(t) \sim (\mathbf{0}, \mathbf{Q}_W)$  as

$$\mathbf{S}_x(\omega) = (i\omega\mathbf{E} - \mathbf{A})^{-1} \mathbf{B} \mathbf{Q}_W \mathbf{B}^T (-i\omega\mathbf{E} - \mathbf{A})^{-T}. \quad (7.5)$$

Thus, the PSD-matrix  $\mathbf{S}_x(\omega)$  is gained from the input intensity matrix  $\mathbf{Q}_W$  by matrix multiplications with the frequency response matrix  $\mathbf{F}_x = (i\omega\mathbf{E} - \mathbf{A})^{-1}$ . Since the state

vector  $\mathbf{x}(t)$  and the system matrix  $\mathbf{A}$  often can be partitioned it is possible to get intermediate results for the PSD-matrices of the road excitation, the suspension system and the chassis, respectively. This reduces possibly the numerical effort considerable. The scalar PSD of the scalar acceleration  $a(t)$ , given by (7.3), reads

$$S_a(\omega) = (\mathbf{c}^T \mathbf{F}_x(\omega) \mathbf{B} + \mathbf{d}^T) \mathbf{Q}_W (\mathbf{B}^T \mathbf{F}_x^T(-\omega) \mathbf{c} + \mathbf{d}), \quad (7.6)$$

where the  $n \times n$ -frequency response matrix  $\mathbf{F}_x$  has been used again. Eq.(7.5) shows that the PSD is a quadratic expression, i. e., the sum of quantities leads to additional terms in the result. Using the PSD (7.6), the frequency weighting required due to (6.10) can be performed very easily. It remains

$$S_{\bar{a}}(\omega) = \alpha^2 |F(\omega)|^2 S_a. \quad (7.7)$$

The variance follows from the PSD by integration over an infinite interval,

$$\sigma_{\bar{a}}^2 = \int_{-\infty}^{\infty} S_{\bar{a}}(\omega) d\omega. \quad (7.8)$$

The numerical computation of the PSD (7.8) is generally not difficult. However, one has to deal with complex matrices, many approximation points and a very large integration interval, this can lead to numerical errors and considerable computation times. The integration (7.8) can be avoided applying the covariance analysis as shown in the next section.

Special care requires the spectral analysis of multi-axle vehicles. The excitation at different axles is given by an excitation function with time delay,

$$\begin{aligned} \mathbf{B}\mathbf{w}(t) &= \sum_{i=1}^m \mathbf{B}_i \zeta_i(t), \\ \zeta_i(t) &= \zeta(t - t_i), \quad 0 = t_1 < t_2 < \dots < t_m, \\ t_i &= \frac{l_i}{v}, \quad i = 1, \dots, m, \end{aligned} \quad (7.9)$$

where  $l_i$  denotes the distance between the front axle and the  $i$ -th axle and  $v = \text{const}$  is the vehicle speed. An example of an excitation with time delay is shown in Figure 21. The PSD-matrix of a two-axle vehicle under random excitation reads

$$\begin{aligned} \mathbf{S}_x(\omega) &= (i\omega \mathbf{E} - \mathbf{A})^{-1} [\mathbf{B}_1 \mathbf{S}_{\zeta} \mathbf{B}_1^T + \mathbf{B}_2 \mathbf{S}_{\zeta} \mathbf{B}_2^T + e^{i\omega(t_1 - t_2)} \mathbf{B}_2 \mathbf{S}_{\zeta} \mathbf{B}_1^T \\ &\quad + e^{i\omega(t_2 - t_1)} \mathbf{B}_1 \mathbf{S}_{\zeta} \mathbf{B}_2^T] (-i\omega \mathbf{E} - \mathbf{A})^{-T}. \end{aligned} \quad (7.10)$$

In frequency domain, the time delay  $(t_2 - t_1)$  results in additional terms, weighted by the exponential function.

### 7.3 Covariance Analysis

In contrast to spectral density analysis, the covariance analysis yields directly the variances that are required for the assessment of the vehicle performance. The covariance

matrix of the entire vehicle system follows from an algebraic equation, the so-called Lyapunov matrix equation, and no integrations are required. An essential prerequisite of the covariance analysis is a white noise input process. This can always be provided by modelling the random vehicle excitation by means of a shape filter.

The Lyapunov matrix equation corresponding to the state (7.1) of the entire vehicle system, reads

$$\mathbf{A}\mathbf{P}_x + \mathbf{P}_x\mathbf{A}^T + \mathbf{B}\mathbf{Q}_W\mathbf{B}^T = \mathbf{0}, \quad (7.11)$$

where  $\mathbf{P}_x = \mathbf{E}\{\mathbf{x}\mathbf{x}^T\}$  denotes the symmetric  $n \times n$ -covariance matrix and  $\mathbf{Q}_W$  is the  $r \times r$ -intensity matrix of the white noise input. An extensive derivation of (7.11) can be found in Müller and Schiehlen (1985). Stable numerical procedures to solve the Lyapunov matrix eq. are given in Smith (1968) and Kreisselmeier (1972). As a result one gets the variances of all state variables of the system. However, the computation of the ride comfort requires the variance of the frequency weighted acceleration. Thus, the covariance analysis has to be applied to the extended system (7.4). The covariance matrix of the extended state vector  $\tilde{\mathbf{x}}$  reads

$$\mathbf{P}_{\tilde{\mathbf{x}}} = \begin{bmatrix} \mathbf{P}_x & \mathbf{P}_{x\bar{v}} \\ \mathbf{P}_{\bar{v}x} & \mathbf{P}_{\bar{v}} \end{bmatrix} \quad (7.12)$$

where  $\mathbf{P}_{x\bar{v}} = \mathbf{P}_{\bar{v}x}^T$  holds. The  $s \times 1$ -vector process  $\bar{\mathbf{v}}(t)$  yields according to (6.15) immediately the frequency weighted acceleration and its variance

$$\sigma_{\bar{a}}^2 = \alpha^2 \bar{\mathbf{h}}^T \mathbf{P}_{\bar{v}} \bar{\mathbf{h}}, \quad (7.13)$$

This is the relation corresponding to (7.8). The investigation of multi-axle vehicles is also possible by the covariance analysis. The theory needed can be developed, as shown in Müller et al. (1980). In contrast to spectral analysis, the covariance analysis can also be applied to nonstationary and nonlinear problems without difficulties.

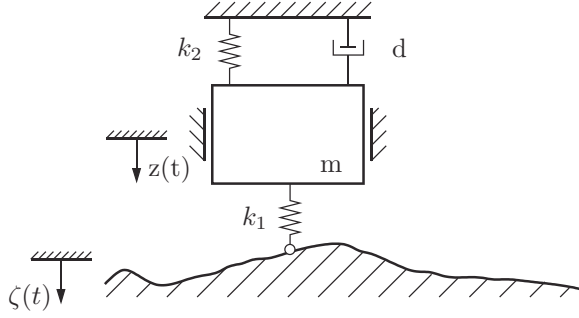
**Random vibrations of a single wheel.** A wheel guided in vertical direction is excited by the road roughness and performs random vibration, see Figure 26. The variance of the dynamic wheel load variations have to be investigated applying covariance analysis. The road roughness is characterized by white velocity noise  $\zeta(t) \sim (0, q)$ . The equation of motion of the wheel reads

$$m\ddot{z} + d\dot{z} + (k_1 + k_2)z = k_1\zeta(t), \quad (7.14)$$

where  $\zeta(t)$  is a scalar random process. After differentiation with respect to time the corresponding state equation has the form

$$\underbrace{\begin{bmatrix} \dot{z} \\ \ddot{z} \end{bmatrix}}_{\tilde{\mathbf{x}}} = \underbrace{\begin{bmatrix} 0 & 1 \\ -\kappa_{12} & -\delta \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{bmatrix} z \\ \dot{z} \end{bmatrix}}_{\tilde{\mathbf{x}}} + \underbrace{\begin{bmatrix} 0 \\ \kappa_1 \end{bmatrix}}_{\mathbf{B}} \dot{\zeta}(t), \quad (7.15)$$

$$\tilde{\mathbf{x}} = \mathbf{A} \tilde{\mathbf{x}} + \mathbf{B} w(t),$$



**Figure 26.** Vehicle wheel subject to random excitation

where the abbreviations  $\delta = d/m$ ,  $\kappa_{12} = (k_1 + k_2)/m$ ,  $\kappa_1 = k_1/m$  have been used and the random excitation is given by a white velocity noise process. Then, the Lyapunov equation reads

$$\mathbf{A}\mathbf{P}_{\dot{x}} + \mathbf{P}_{\dot{x}}\mathbf{A}^T + \mathbf{B}\mathbf{q}\mathbf{B}^T = 0, \quad (7.16)$$

where the  $2 \times 2$ -covariance matrix of the first derivative  $\dot{x}$  of the state vector has been introduced,

$$\mathbf{P}_{\dot{x}} = \begin{bmatrix} P_{11} & P_{12} \\ P_{12} & P_{22} \end{bmatrix} = \begin{bmatrix} P_{z\dot{z}} & P_{z\ddot{z}} \\ P_{z\dot{z}} & P_{\dot{z}\ddot{z}} \end{bmatrix}. \quad (7.17)$$

The solution of (7.16) yields the following matrices

$$\mathbf{A}\mathbf{P}_{\dot{x}} = \begin{bmatrix} P_{12} & P_{22} \\ -\kappa_{12}P_{12} - \delta P_{12} & -\kappa_{12}P_{12} - \delta P_{22} \end{bmatrix}, \quad (7.18)$$

$$\mathbf{B}\mathbf{q}\mathbf{B}^T = \begin{bmatrix} 0 & 0 \\ 0 & \kappa_1^2 q \end{bmatrix}. \quad (7.19)$$

Due to the symmetry of the covariance matrix from (7.16) it follows altogether three linear equations for the unknowns

$$\begin{bmatrix} 0 & 2 & 0 \\ -\kappa_{12} & -\delta & 1 \\ 0 & 2\kappa_{12} & 2\delta \end{bmatrix} \begin{bmatrix} P_{11} \\ P_{12} \\ P_{22} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \kappa_1^2 q \end{bmatrix} \quad (7.20)$$

having the solutions

$$P_{11} = \frac{k_1^2}{2d(k_1 + k_2)}q, \quad P_{12} = 0, \quad P_{22} = \frac{k_1^2}{2dm}q. \quad (7.21)$$

The weighting vector  $\mathbf{b}$  of the dynamic wheel loads reads according to (1)

$$f_{Dym} = k_1(z - \zeta) = \underbrace{\begin{bmatrix} -d & -m \end{bmatrix}}_{\mathbf{b}^T} \underbrace{\begin{bmatrix} \dot{z} \\ \ddot{z} \end{bmatrix}}_{\dot{\mathbf{x}}}, \quad (7.22)$$

where  $k_2 \ll k_1$  has been regarded. The variance  $\sigma_f^2$  of the dynamic wheel load variation reads

$$\sigma_f^2 = \mathbf{b}^T \mathbf{P}_{\dot{\mathbf{x}}} \mathbf{b} = \left[ \frac{k_1 d}{2} + \frac{k_1^2 m}{2d} \right] q. \quad (7.23)$$

It can be seen, that  $\sigma_f^2$  becomes a minimum if the viscous damping factor  $D = d/(2\sqrt{k_1 m}) = 0.5$  is chosen. In this case of optimal damping it follows  $\sigma_f^2 = k_1 \sqrt{k_1 m} q$ . This result shows that small wheel masses  $m$  and small tire stiffnesses  $k_1$  are beneficial for the dynamic wheel load variations. However, there are design limits for these parameters.

The assumption  $k_2 \ll k_1$  is necessary in this example, since an idealized white velocity noise process has been considered. In the general case of coloured noise this assumption is not required.

## 8 Suspension Analysis

For the design of vehicle suspensions the random road unevenness profiles have to be taken into account. There are two main tasks of a vehicle suspension. On the one hand the vehicle body has to be carried in constant height without vibrational accelerations while on the other hand the vehicle wheels should follow the uneven profile of the road without any delay resulting in a constant wheel load. Both requirements are inconsistent with each other, and, therefore, the problem may be solved by frequency decoupling. Then, the eigenfrequency of the body should be as low as possible while the eigenfrequency of the wheels has to be as high as possible. However, there are design constraints which restrain the frequency decoupling, like the limited relative motion space between of the wheels and the body. This means that an optimal design has to consider the engineering constraints, and the design process has to include the vehicle structure and the randomly uneven road.

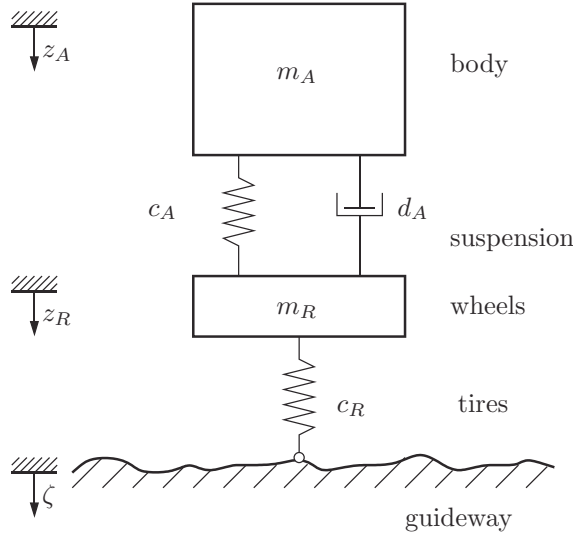
### 8.1 Random Vibrations of a Quarter Car Model

The fundamentals of vehicle suspensions are presented for a quarter car model, Figure 27, which offers frequency decoupling by two degrees of freedom and includes random excitation by road unevenness. With respect to the equilibrium condition the linear equations of motion read as

$$\begin{aligned} & \begin{bmatrix} m_A & 0 \\ 0 & m_R \end{bmatrix} \begin{bmatrix} \ddot{z}_A \\ \ddot{z}_R \end{bmatrix} + \begin{bmatrix} d_A & -d_A \\ -d_A & d_A \end{bmatrix} \begin{bmatrix} \dot{z}_A \\ \dot{z}_R \end{bmatrix} \\ & + \begin{bmatrix} c_A & -c_A \\ -c_A & c_A + c_R \end{bmatrix} \begin{bmatrix} z_A \\ z_R \end{bmatrix} = \begin{bmatrix} 0 \\ c_R \end{bmatrix} \zeta(t) \end{aligned} \quad (8.1)$$

where  $m_A$  is the body mass,  $m_R$  is the generalized mass of all wheels as well as  $d_A$ ,  $c_A$  and  $c_R$  represent the generalized damping coefficient, body spring coefficient and tire spring coefficient of all wheels. As excitation a white velocity noise with intensity  $q$  is assumed,

$$\dot{\zeta}(t) \sim (0, q). \quad (8.2)$$



**Figure 27.** Vehicle model for the analysis of vertical vibrations

Using the abbreviations

$$a = \frac{c_A}{m_A}, \quad b = \frac{c_A}{m_R}, \quad c = \frac{c_R}{m_R}, \quad d = \frac{d_A}{m_A}, \quad e = \frac{d_A}{m_R} \quad (8.3)$$

one gets from (8.1) the state equations which have been differentiated once to obtain an velocity excitation of the vehicle,

$$\underbrace{\begin{bmatrix} \ddot{z}_A \\ \ddot{z}_R \\ \dot{z}_A \\ \dot{z}_R \end{bmatrix}}_{\dot{\mathbf{x}}} = \underbrace{\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -a & a & -d & d \\ b & -b-c & e & -e \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{bmatrix} \dot{z}_A \\ \dot{z}_R \\ z_A \\ z_R \end{bmatrix}}_{\mathbf{x}} + \underbrace{\begin{bmatrix} 0 \\ 0 \\ 0 \\ c \end{bmatrix}}_{\mathbf{B}} \underbrace{\dot{\zeta}(t)}_{\mathbf{w}(t)} \quad (8.4)$$

For the assessment of the vehicles performance driving comfort and driving safety are used. It yields for the vertical acceleration

$$a_A = \ddot{z}_A = \begin{bmatrix} 0 & 0 & 1 & 0 \end{bmatrix} \dot{\mathbf{x}} = \mathbf{c}^T \dot{\mathbf{x}} \quad (8.5)$$

and for the dynamical tire load

$$f = c_R(\zeta - z_R) = m_A \ddot{z}_A + m_R \ddot{z}_R = \begin{bmatrix} 0 & 0 & m_A & m_R \end{bmatrix} \dot{\mathbf{x}} = \mathbf{e}^T \dot{\mathbf{x}} \quad (8.6)$$

where the  $4 \times 1$ -weighting vectors are introduced. Then, it remains for the standard deviations of non-weighted acceleration and wheel load

$$\sigma_a^2 = \mathbf{c}^T \mathbf{P} \mathbf{c} = P_{33} \quad (8.7)$$

and

$$\sigma_f^2 = \mathbf{e}^T \mathbf{P} \mathbf{e} = m_A^2 P_{33} + 2m_A m_R P_{34} + m_R^2 P_{44} \quad (8.8)$$

with the stationary  $4 \times 4$ -covariance matrix

$$\mathbf{P} = \mathbb{E}\{\dot{\mathbf{x}}(t)\dot{\mathbf{x}}^T(t)\} = \text{const} . \quad (8.9)$$

The problem is reduced now to the evaluation of the covariance matrix (8.9) of the system (8.4). For this purpose the covariance analysis is most adequate offering analytical solutions for low order systems in contrary to the power spectral density approach. An extensive presentation of the covariance analysis can be found in Müller et al. (1980). The corresponding Lyapunov matrix equation reads

$$\mathbf{A}\mathbf{P} + \mathbf{P}\mathbf{A}^T + \mathbf{B}\mathbf{q}\mathbf{B}^T = 0 . \quad (8.10)$$

For the solution a matrix polynomial is applied as outlined by Müller and Schiehlen (1985). This polynomial reads for  $n = 4$  as follows

$$\mathbf{P} = \frac{1}{2 \det \mathbf{H}} \sum_{k=0}^3 H_{k+1,1} \sum_{m=0}^{2k} (-1)^m \mathbf{A}_m \mathbf{Q} \mathbf{A}_{2k-m}^T \quad (8.11)$$

where

$$\mathbf{H} = \begin{bmatrix} a_1 & 1 & 0 & 0 \\ a_3 & a_2 & a_1 & 1 \\ 0 & a_4 & a_3 & a_2 \\ 0 & 0 & 0 & a_4 \end{bmatrix} \quad (8.12)$$

is the  $4 \times 4$ -Hurwitz matrix also known from stability theory and  $H_{k+1,1}$  are the corresponding scalar cofactors. Further,

$$a_i, \quad i = 1(1)4 \quad (8.13)$$

mean the characteristic coefficients of the system matrix  $\mathbf{A}$ . In addition, the  $4 \times 4$  - auxiliary matrices are defined by

$$\mathbf{A}_m = \mathbf{A}\mathbf{A}_{m-1} + a_m \mathbf{E}, \quad m = 0(1)6 \quad (8.14)$$

where it yields  $\mathbf{A}_p = \mathbf{0}$ ,  $p = 4(1)6$ . The  $4 \times 4$ -intensity matrix reads as

$$\mathbf{Q} = \mathbf{q}\mathbf{B}\mathbf{B}^T . \quad (8.15)$$

Some intermediate results are listed for the following quantities. Characteristic equation:

$$\det(\lambda \mathbf{E} - \mathbf{A}) = \lambda^4 + (d + e)\lambda^3 + (a + b + c)\lambda^2 + cd\lambda + ac = 0 . \quad (8.16)$$

Determinant of Hurwitz matrix:

$$\det \mathbf{H} = a_4 H_3 = ac(c^2 de) . \quad (8.17)$$



Cofactors of Hurwitz matrix:

$$\left. \begin{aligned} H_{11} &= a_4(a_2a_3 - a_1a_4) = ac^3d, \\ H_{21} &= -a_3a_4 = -ac^2d, \\ H_{31} &= a_1a_4 = ac(d+e), \\ H_{41} &= -(a_1a_2 - a_3) = -(a+b)(d+e) - ce. \end{aligned} \right\} \quad (8.18)$$

Auxiliary matrices multiplied by  $4 \times 1$ -matrix  $\mathbf{B}$ :

$$\begin{aligned} \mathbf{A}_0\mathbf{B} &= \mathbf{B} = c[0 \ 0 \ 0 \ 1]^T, \\ \mathbf{A}_1\mathbf{B} &= \mathbf{A}\mathbf{B} + a_1\mathbf{B} = c[0 \ 1 \ d \ d]^T, \\ \mathbf{A}_2\mathbf{B} &= \mathbf{A}\mathbf{A}_1\mathbf{B} + a_2\mathbf{B} = c[d \ d \ a \ a]^T, \\ \mathbf{A}_3\mathbf{B} &= \mathbf{A}\mathbf{A}_2\mathbf{B} + a_3\mathbf{B} = c[a \ a \ 0 \ 0]^T, \\ \mathbf{A}_4\mathbf{B} &= \mathbf{0}. \end{aligned} \quad (8.19)$$

Then, the elements of the covariance matrix required for the assessments (8.7) and (8.8) are easily found as

$$P_{33} = \frac{q}{2de} [cd^3 + a^2(d+e)], \quad (8.20)$$

$$P_{34} = \frac{q}{2de} [cd(d^2 - a) + a^2(d+e)], \quad (8.21)$$

$$P_{44} = \frac{q}{2de} [(a-c)^2d + cd^3 + a^2e]. \quad (8.22)$$

Inserting (8.20) to (8.22) in (8.7) and (8.8), and eliminating the abbreviations (8.3), it remain the final results

$$\sigma_a^2 = \frac{q}{2} \left[ \frac{c_R d_A}{m_A^2} + \frac{c_A^2 (m_A + m_R)}{d_A m_A^2} \right], \quad (8.23)$$

$$\begin{aligned} \sigma_f^2 &= \frac{q}{2} \left[ \left(1 + \frac{m_R}{m_A}\right)^3 \frac{c_A^2 m_A}{d_A} + \left(1 + \frac{m_R}{m_A}\right)^2 c_R d_A \right. \\ &\quad \left. - 2 \left(1 + \frac{m_R}{m_A}\right) \frac{c_A c_R m_R}{d_A} + \frac{c_R^2 m_R}{d_A} \right]. \end{aligned} \quad (8.24)$$

Thus, analytical explicit solutions are found which can be used for optimization of the original design parameters.

The optimization will be performed for a compact passenger car with a body mass  $m_A = 1200$  kg. Then, there remain four design variables with the following reference values:

Wheel mass	$m_R = 80$ kg ,
Body spring	$c_A = 30000$ N/m ,
Tire spring	$c_R = 320000$ N/m ,
Shock absorber	$d_A = 4800$ Ns/m .

These design variables are generalized quantities. E.g., the wheel mass  $m_R$  includes not only the masses of the four wheels but also the corresponding moments of inertia of the control arms of the wheel suspensions.

For the discussion of the sensitivity with respect to the design variables the following specific quantities are useful.

$$\frac{\sigma_a^2(c_A, d_A)}{\sigma_{a\ reference}^2} \quad \text{Driving comfort by suspension strut}$$

$$\frac{\sigma_f^2(c_A, d_A)}{\sigma_{f\ reference}^2} \quad \text{Driving safety by suspension strut}$$

$$\frac{\sigma_a^2(c_R, m_R)}{\sigma_{a\ reference}^2} \quad \text{Driving comfort by axle and tire}$$

$$\frac{\sigma_f^2(c_R, m_R)}{\sigma_{f\ reference}^2} \quad \text{Driving safety by axle and tire}$$

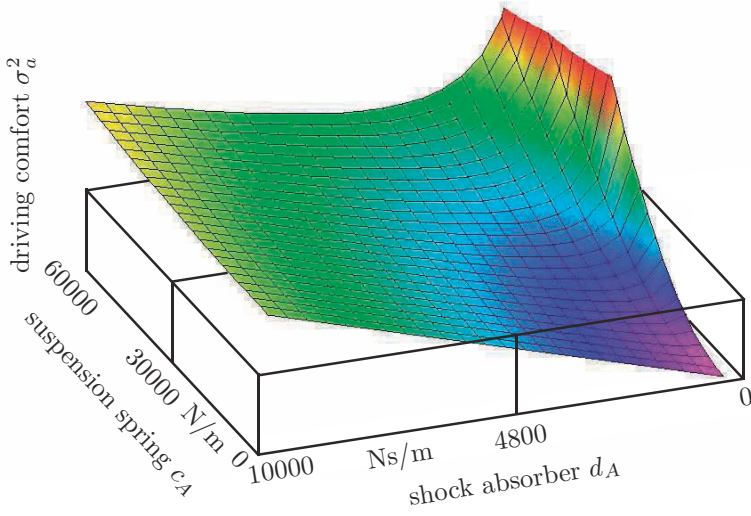
Thus, the original four-dimensional optimization problem is reduced to two two-dimensional optimization problems which can be treated graphically very nicely, Figure 28 to Figure 31, where the height of the box represents driving comfort and driving safety, respectively, of the reference values. For the strut parameters the following tendencies are found.

1. Large and small damper coefficients are harmful, for medium damping comfort and safety are optimal.
2. Small spring and damper coefficients improve the comfort considerably. However, soft springs result in large variations of the equilibrium position due to body mass variations which may only be compensated by a ride-height control system.
3. The safety is independent of the spring coefficient.
4. The reference value of the damping coefficient yields optimal safety. The damping coefficient with optimal comfort deteriorates the safety essentially. Since the efficiency of a damper decreases with time, drivers feel in used cars more comfortable forgetting about safety.

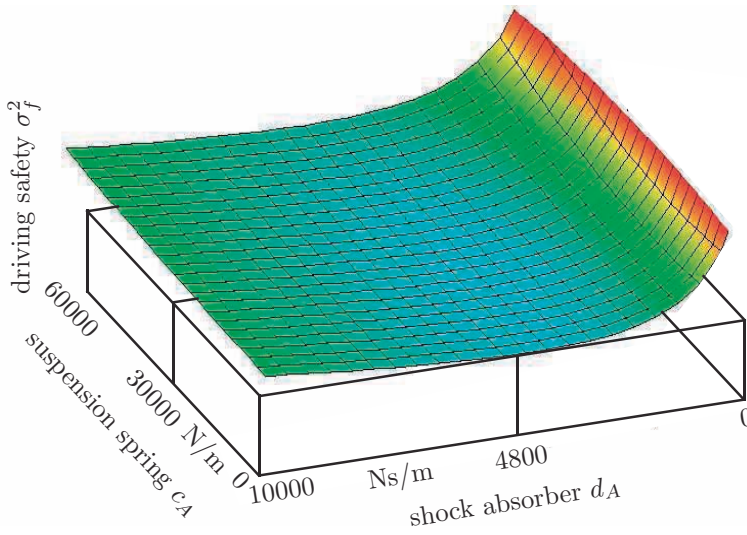
For axle and tire parameters also some general tendencies are observed.

1. The comfort is independent of the tire mass.
2. Small tire mass improves safety essentially. The generalized mass can be also reduced by a suitable suspension design. Nevertheless, the reductions are limited.
3. Small tire stiffness improves comfort and safety. The air pressure in the tires is responsible for the tire stiffness and can not be reduced essentially. Low pressure results in strong wear of the tire and weak lateral guidance of the vehicle. Therefore, the reference value of the tire spring coefficient is a boundary optimum.

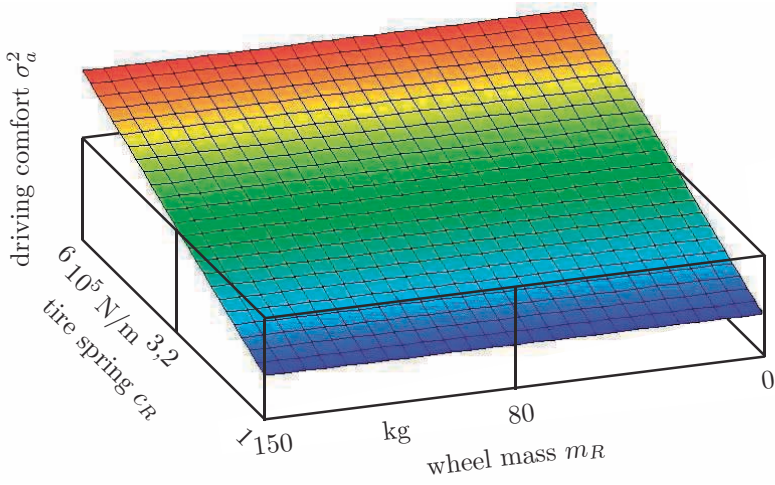
In summary, the reference values of the suspension are close to an optimum which was found empirically during the development of the automobile in more than one century. Further progress will be achieved in the future by mechatronic system design where the stochastic methods presented will be most helpful.



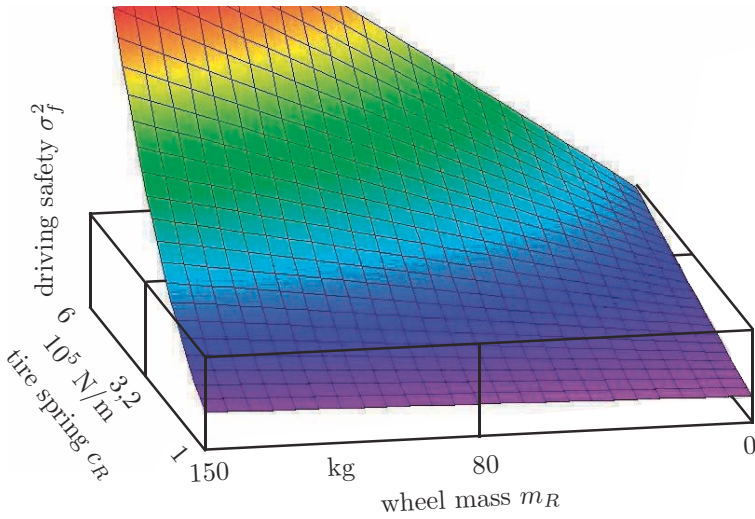
**Figure 28.** Driving comfort by suspension strut design variables



**Figure 29.** Driving safety by suspension strut design variables



**Figure 30.** Driving comfort by axle and tire design variable

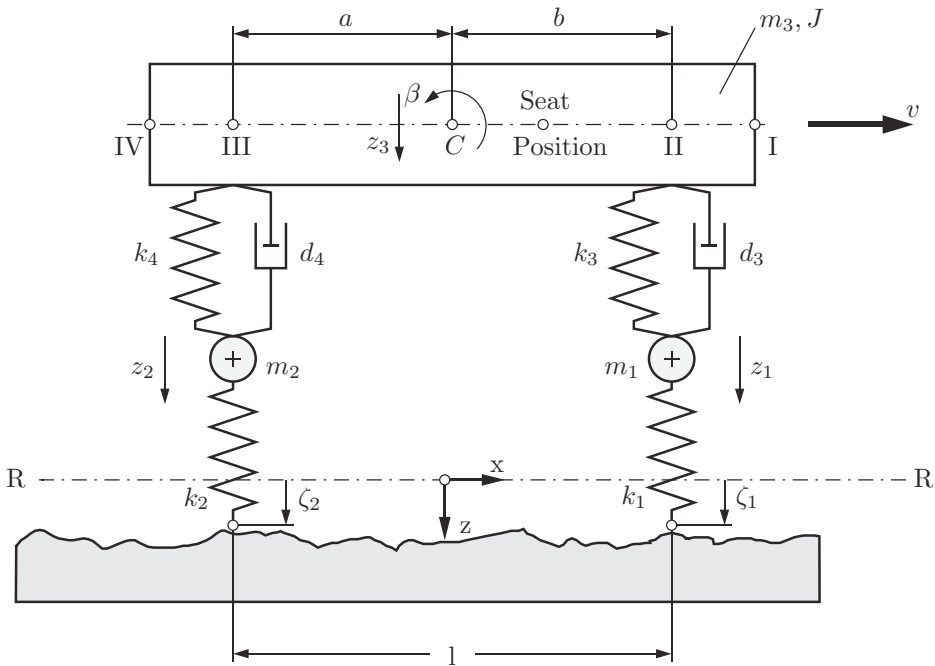


**Figure 31.** Driving safety by axle and tire design variable

In addition to the qualitative results reported in this paper, more quantitative results can be obtained for specific vehicles with more complex nonlinear models represented by the equations of motion (3.1) and coloured noise excitation (5.42). However, for the fundamental understanding of the principles of driving comfort and driving safety a linear model with white velocity noise is most adequate.

## 8.2 Random vibrations of a two axle vehicle.

The covariance analysis is now demonstrated for the two axle vehicle shown in Figure 32. A two axle vehicle with four degrees of freedom is excited by white noise (5.51) for the derivated process  $\dot{\zeta}(t)$ .



**Figure 32.** Two axle vehicle model traveling on rough road

The vehicle parameters read as

$m_1$	=	100 kg	$k_4$	=	51208 N/m
$m_2$	=	54 kg	$d_3$	=	3198 Ns/m
$m_3$	=	1247 kg	$d_4$	=	3434 Ns/m
$J$	=	1945 kgm <sup>2</sup>	$a$	=	1.2 m
$k_1$	=	$k_2 = 343350$ N/m	$b$	=	1.3 m
$k_3$	=	47382 N/m	$c$	=	2.5 m

Small motions are assumed and the time lag  $t_2 = l_2/v$  between front and rear contact has to be taken into account.

The generalized coordinates are summarized in the position vector  $\mathbf{y}(t)$ ,  $\mathbf{y}(t) = [z_1(t), z_2(t), z_3(t), \beta(t)]^T$ . Then, the equations of motion read

$$\mathbf{M}\ddot{\mathbf{y}}(t) + \mathbf{D}\dot{\mathbf{y}}(t) + \mathbf{K}\mathbf{y}(t) = \mathbf{s}_1\zeta_1(t) + \mathbf{s}_2\zeta_2(t), \quad \zeta_1(t) = \zeta(t), \zeta_2(t) = \zeta(t - t_2), \quad (8.25)$$

where  $\mathbf{M}$ ,  $\mathbf{D}$ ,  $\mathbf{K}$  characterize inertia, damping and stiffness matrix, respectively, and  $\mathbf{s}_1$ ,  $\mathbf{s}_2$  are the excitation input vectors,  $\mathbf{M} = \mathbf{diag}[m_1, m_2, m_3, J]$ ,  $\mathbf{s}_1 = [k_1 \ 0 \ 0 \ 0]^T$ ,  $\mathbf{s}_2 = [0 \ k_2 \ 0 \ 0]^T$ ,

$$\mathbf{D} = \begin{bmatrix} d_3 & 0 & -d_3 & bd_3 \\ 0 & d_4 & -d_4 & -ad_4 \\ -d_3 & -d_4 & d_3 + d_4 & -bd_3 + ad_4 \\ bd_3 & -ad_4 & -bd_3 + ad_4 & b^2d_3 + a^2d_4 \end{bmatrix}, \quad (8.26)$$

$$\mathbf{K} = \begin{bmatrix} k_1 + k_3 & 0 & -k_3 & bk_3 \\ 0 & k_2 + k_4 & -k_4 & -ak_4 \\ -k_3 & -k_4 & k_3 + k_4 & -bk_3 + ak_4 \\ bk_3 & -ak_4 & -bk_3 + ak_4 & b^2k_3 + a^2k_4 \end{bmatrix}. \quad (8.27)$$

From (8.25) the vehicle state space representation can be obtained,

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{b}_1\zeta_1(t) + \mathbf{b}_2\zeta_2(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\boldsymbol{\xi}, \quad (8.28)$$

$$\mathbf{x}(t) = \begin{bmatrix} \mathbf{y}(t) \\ \dot{\mathbf{y}}(t) \end{bmatrix}, \quad \boldsymbol{\xi}(t) = \begin{bmatrix} \zeta_1(t) \\ \zeta_2(t) \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} \mathbf{0} & \mathbf{E} \\ -\mathbf{M}^{-1}\mathbf{K} & -\mathbf{M}^{-1}\mathbf{D} \end{bmatrix}, \quad (8.29)$$

$$\mathbf{b}_1 = \begin{bmatrix} \mathbf{0} \\ \mathbf{M}^{-1}\mathbf{s}_2 \end{bmatrix}, \quad \mathbf{B} = [\mathbf{b}_1 \ \mathbf{b}_2].$$

In order to use white noise vehicle excitation,  $\dot{\zeta}(t) \equiv w(t)$ , (8.28) has to be differentiated,

$$\ddot{\mathbf{x}}(t) = \mathbf{A}\dot{\mathbf{x}}(t) + \mathbf{B}w(t). \quad (8.30)$$

Here, the autocorrelation matrix of the white noise vector process  $\mathbf{w}$  reads

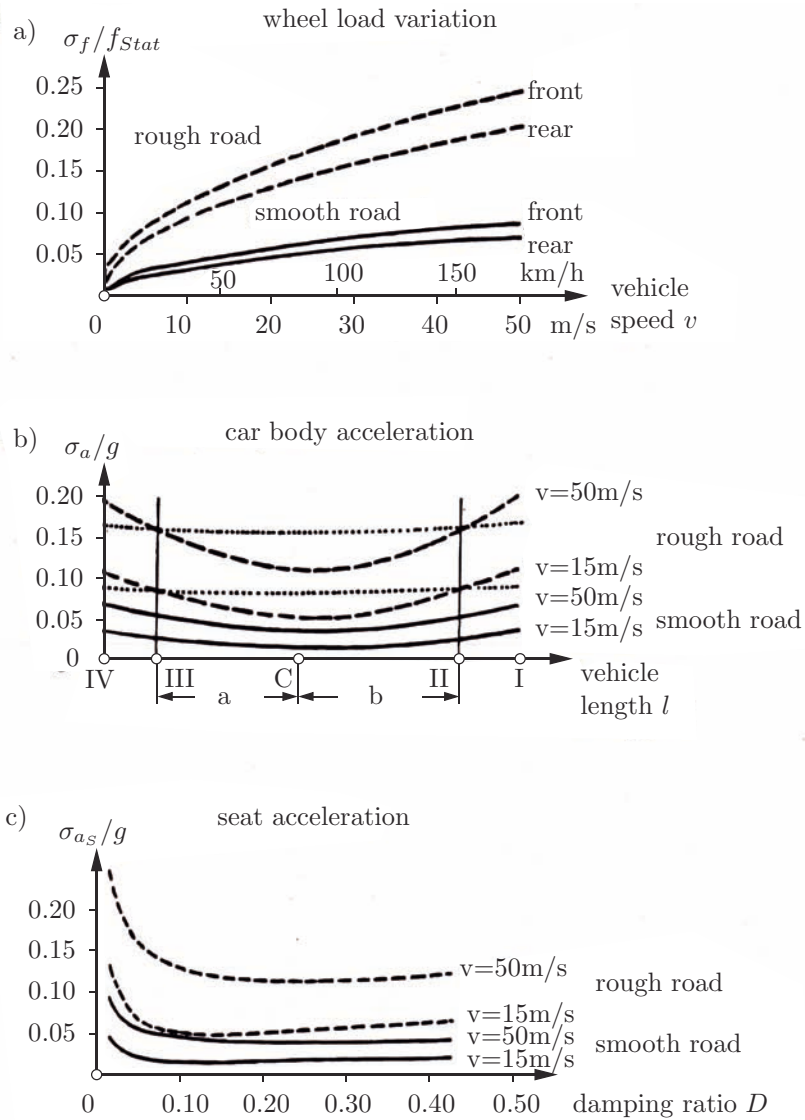
$$\mathbf{R}_w(\tau) = \mathbb{E}\{\mathbf{w}(t)\mathbf{w}^T(t - \tau)\} = q_\zeta \begin{bmatrix} \delta(\tau) & \delta(\tau + t_2) \\ \delta(\tau - t_2) & \delta(\tau) \end{bmatrix}. \quad (8.31)$$

Thus, the covariance matrix  $\mathbf{P}_{\dot{\mathbf{x}}}$  in steady state follows from the extended algebraic Lyapunov equation

$$\mathbf{A}\mathbf{P}_{\dot{\mathbf{x}}} + \mathbf{P}_{\dot{\mathbf{x}}}\mathbf{A}^T + \mathbf{Q} = \mathbf{0}, \quad \mathbf{Q} = q_\zeta \left[ \mathbf{b}_1\mathbf{b}_1^T + \mathbf{b}_2\mathbf{b}_2^T + e^{A t_2}\mathbf{b}_1\mathbf{b}_2^T + \mathbf{b}_2\mathbf{b}_1^T e^{A^T t_2} \right]. \quad (8.32)$$

From  $\mathbf{P}_{\dot{\mathbf{x}}}$  the variances  $\sigma_a^2$  of the car body acceleration  $\ddot{z}_p(t)$  at any position  $P$  with coordinate  $x_p(t) = \mathbf{a}^T\mathbf{y}(t)$  can immediately be calculated by

$$\sigma_a^2 = [\mathbf{0}^T \ \mathbf{a}^T]\mathbf{P}_{\dot{\mathbf{x}}} \begin{bmatrix} \mathbf{0} \\ \mathbf{a} \end{bmatrix}. \quad (8.33)$$



**Figure 33.** Wheel load variation and vertical accelerations of the two axle vehicle

A similar expression holds for the variances  $\sigma^2$  of the wheel loads  $f_i = k_i(\zeta_i - y_i)$ ,  $i = 1, 2$ , but here (8.25) has to be used to get the displacements. Figure 33 shows some numerical results, where a smooth road excitation ( $q_{\dot{\zeta}} = 3,14 \cdot 10^{-6} \cdot v$  m<sup>2</sup>/s, solid line) and a rough road excitation ( $q_{\dot{\zeta}} = 24,7 \cdot 10^{-6} \cdot v$  m<sup>2</sup>/s, dashed line) is assumed.

In Figure 33a) the standard deviation  $\sigma_f/f_{Stat}$  of the load variation of front and rear wheel is plotted against speed  $v$  where  $f_{Stat}$  denotes the static load. It increases with  $\sqrt{v}$  since the excitation intensity is proportional to  $v$ . Even in the worst case (rough road, front wheel, maximum speed) the safety margin  $R$  according to (6.22) takes the value  $R = 1 - \sigma_F/f_{Stat} \approx 0.75$  which indicates a safe ride.

In Figure 33b) the standard deviation  $\sigma_a/g$  of the car body acceleration versus vehicle length is given, which clearly shows a minimum for the center of mass position at different speeds. Neglecting the time lag between front and rear excitation (dotted line) can lead to incorrect results especially near the seat position.

The covariance analysis has also an important application in optimization problems. Figure 33c) shows a simple example, where the influence of the nondimensional damping ratio  $D = D_i \equiv d_i/(2\sqrt{k_i\bar{m}_i})$ ,  $i = 1, 2$ , ( $\bar{m}_1 = m_3a/l_1$ ,  $\bar{m}_2 = m_3b/l_1$ ) on the standard deviation of the seat acceleration,  $\sigma_{a_s}/g$ , is studied. Best results are obtained in the range  $0.1 \leq D \leq 0.25$ . However, tradeoffs with the wheel load variation call for higher damping. For a ride comfort evaluation the perception measure  $K$  has to be calculated. A rough approximate estimation using the thumb rule  $a_{\max} \leq 0.5$  m/s<sup>2</sup> shows that the worst case (rough road, maximum speed, maximum damping) certainly results in a bad ride comfort.

## Bibliography

- VDI Richtlinie 2057. *Beurteilung der Einwirkung mechanischer Schwingungen auf den Menschen*. Düsseldorf, 1975-1979.
- ISO International Standard 2631. *Guide for the Human Exposure to Whole-Body Vibrations*, 1974.
- D.S. Bae and E.J. Haug. A recursive formulation for constrained mechanical system dynamics: Part i, open loop systems. *Mechanics of Structures and Machines*, 15: 359–382, 1987a.
- D.S. Bae and E.J. Haug. A recursive formulation for constrained mechanical system dynamics: Part ii, closed loop systems. *Mechanics of Structures and Machines*, 15: 481–506, 1987b.
- V. Bormann. Messungen von Fahrbahnunebenheiten paralleler Fahrspuren und Anwendung der Ergebnisse. *Vehicle System Dynamics*, 7, 1978.
- H. Brandl, R. Johanni, and M. Otter. A very efficient algorithm for the simulation of robots and similar multibody systems without inversion of the mass matrix. In P. Kopacek, I. Troch, and K. Desoyer, editors, *Theory of Robots*, pages 95–100. Pergamon, Oxford, 1988.
- H. Braun. *Untersuchung von Fahrbahnunebenheiten und Anwendung der Ergebnisse*. Dr.-Ing. Diss, Braunschweig, 1969.
- S.H. Crandall and W.D. Mark. *Random Vibration in Mechanical Systems*. Academic Press, New York, London, 1963.



- L. Czerny. Analyse stationärer Zufallsschwingungen. In *Fort. Ber. VDI-Reihe 11 Nr. 99*, Düsseldorf, 1987. VDI-Verlag.
- C.J. Dodds and J.D. Robson. The description of road surface roughness. *J. Sound Vibrations*, 31:175–183, 1973.
- W. Heinrich and K. Hennig. *Zufallsschwingungen mechanischer Systeme*. Vieweg, Braunschweig, 1978.
- J.M. Hollerbach. A recursive lagrangian formulation of manipulator dynamics and comparative study of dynamics formulation complexity. *IEEE Trans. Syst. Man. Cybern.*, 11:730–736, 1980.
- T.R. Kane and D.A. Levinson. *Dynamics: Theory and Applications*. McGraw Hill, New York, 1985.
- G. Kreisselmeier. A solution of the bilinear matrix equation  $AY + YB = -Q$ . *SIAM J. Appl. Math.*, 23:334–338, 1972.
- E. Kreuzer and G. Leister. Programmsystem NEWEUL'90. In *Anleitung AN-24*, Stuttgart, 1991. Institut für Technische und Numerische Mechanik.
- K. Magnus and H.H. Müller. *Grundlagen der Technischen Mechanik*. Teubner, Wiesbaden, 1990.
- M. Mitschke. *Dynamik der Fahrzeuge*. Springer, Berlin, 1972.
- P.C. Müller and K. Popp. Kovarianzanalyse von linearen Zufallsschwingungen mit zeitlich verschobenen Erregerprozessen. *Z. angew. Math. Mech.*, 59:T144–T146, 1979.
- P.C. Müller and W. Schiehlen. *Linear Vibrations*. Martinus Nijhoff, Dordrecht, 1985.
- P.C. Müller, K. Popp, and W.O. Schiehlen. Berechnungsverfahren für stochastische Fahrzeugschwingungen. *Ing. Arch.*, 49:235–254, 1980.
- D.E. Newland. *Random Vibrations and Spectral Analysis*. Longmann, New York, London, 1975.
- G. Rill. *Instationäre Fahrzeugschwingungen bei stochastischer Erregung*. Dr.-Ing Diss, Stuttgart, 1983.
- G. Rill. Auswahl eines geeigneten Integrationsverfahren für nichtlineare Bewegungsgleichungen bei Erregung durch beliebige Zeitfunktionen. In *Forschungsbericht FB-4*, Stuttgart, 1981. Institut für Technische und Numerische Mechanik.
- S.K. Saha and W. Schiehlen. Recursive kinematics and dynamics for parallel structural closed-loop multibody systems. *Mechanics of Structures and Machines*, 29:143–175, 2001.
- W. Schiehlen, editor. *Multibody System Handbook*. Springer, Berlin, 1990.
- W. Schiehlen. Computational aspects in multibody system dynamics. *Comp. Meth. Appl. Mech. Eng.*, 90:569–582, 1991.
- W. Schiehlen and P. Eberhard. *Technische Dynamik*. Teubner, Wiesbaden, 2004.
- L.F. Shampine and M.K. Gordon. *Computer Solution of Ordinary Differential Equations*. Freeman, San Francisco, 1984. Vieweg, Braunschweig, 1984.
- R.A. Smith. Matrix equation  $XA + BX = C$ . *SIAM J. Appl. Math.*, 16:198–201, 1968.
- C. Voy. Die Simulation vertikaler Fahrzeugschwingungen. In *Fort. Ber. VDI Reihe 12 Nr.30*, Düsseldorf, 1978. VDI-Verlag.

# Dynamics of Railway Vehicles and Rail/Wheel Contact

Hans True \*

Informatics and Mathematical Modelling, The Technical University of Denmark,  
Kgs.Lyngby, Denmark

**Abstract** In these notes the fundamentals of the mechanics of rail/wheel contact and deterministic vehicle dynamics is explained. Chapter 1 describes the kinematics and dynamics of rail/wheel contact. Chapter 2 explains why vehicle dynamics must be treated as a nonlinear dynamic problem and how the model problem must be formulated. Chapters 3 and 4 deal with the theory of nonlinear parameter dependent dynamic systems in general, and chapter 5 yields the proof that the theory also applies to the high-dimensional vehicle dynamic problems. In chapter 6 the concept of non-smooth systems is introduced and some of the most important dynamic effects on vehicle dynamic problems are described. In chapter 7 characteristic features of railway vehicle dynamics are described, and in chapter 8 recommendations are presented for the numerical handling that is necessary for the investigation of vehicle dynamic problems. The notes end with a guide to vehicle system dynamics and 47 literature references.

## 1 Wheel/Rail Contact

### 1.1 Kinematics

The fundamental guidance system of railways consists of a flanged wheel set rolling on a track with two rails. A clearance between the outer edge of the flange of the wheel and the inner face of the rail is provided to prevent squeezing of the wheels between the rails. A plane through the wheelset containing its centre line will intersect the wheel tread in a curve called *the wheel profile*. The angle  $\delta$  between the tangent to the wheel profile and the centre line of the wheel set measured in radians is called *the contact angle*. It is made positive in most points on the wheel tread. The positive contact angle makes the wheel set kinematically self-centering. Negative contact angles that may develop as a result of wear make the motion of the wheel set unstable and should be avoided.

As long as the wheel set rolls along a straight track, the flanges ought not contact the rails. Disturbances in the track geometry will push the wheel set away from its centred position but owing to the stabilizing effect of the wheel profile the wheel set will then oscillate horizontally around the centre line of the track. Flange contact will only occur

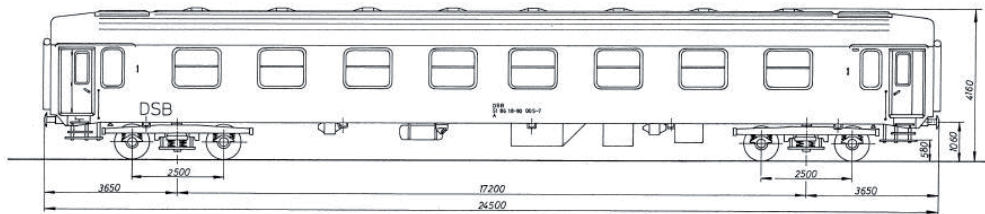
---

\*I thank Railnet Denmark and Banverket, Sweden, for their financial support of the preparation of these lecture notes.

if the disturbances are sufficiently large. When it occurs, the restoring force suddenly grows very fast.

Problems will obviously arise, when a vehicle with a fixed wheel base must negotiate a curve. The longer the wheel base the smaller the curvature must be in order to prevent flange contact and squeezing. On the other hand it is also obvious that the longer the fixed wheel base is, the smaller will the yaw motion be of a vehicle running on a straight but imperfect track. The longer wheel base therefore leads to an improvement in riding quality on a straight track.

In an effort to solve the conflict between good curving behaviour and riding quality the railway companies introduced rail vehicles with two or more wheel sets in under-carriages - one under-carriage under each end of the long vehicle - instead of the single wheel sets (see Figure 1). In North America, where they were first introduced in the middle of the nineteenth century, the under-carriages are called *trucks*. In Europe they are called *bogies* (bo<sup>u</sup>giz). The car body of bogie vehicles is supported by the frame of the bogies, and the bogie frame is supported by the wheel sets. Systems of springs and dampers are built into the supports to improve the riding quality in the car body. The support system between the wheel sets and the bogie frame is called *the primary suspension* and the support system between the frame and the car body is called *the secondary suspension*. Freight wagons often have only one support system. The American three-piece-freight truck has only a secondary suspension. The European standard two-axle freight wagon has no bogies and therefore only one primary suspension system.



**Figure 1.** A four-axle bogie passenger car.

The radius of the wheels decrease from the flange to the field side - at least when they are new. The stabilizing effect of the wheel profile is due to the difference in rolling radii, when the wheel set runs off centre. When one of the wheels of a wheel set runs on a larger radius than the other wheel, then the wheel set will turn back towards the equilibrium position where the rolling radii are equal. Without damping, i.e. in the purely kinematic situation, the wheel set will overshoot the equilibrium position and the wheel set will start yaw oscillations. The important influence of this so-called 'effective conicity' has been recognized for some time. Notice that the 'effective conicity' depends on the lateral and yaw positions of the wheel set, and it varies with the displacements. The stabilizing action of the wheel profile is the cause of a kinematic instability, which was treated by Klingel (1883) in a celebrated paper. The oscillation that results from

this kinematic instability is *neutrally stable* or *Lyapunov stable*. These concepts will be defined later.

The Klingel instability is a kinematic product, and it must not be confused with the so-called *hunting*, which is an asymptotically stable oscillating motion that is a product of the dynamics. In general, however, the 'effective conicity' has a destabilizing effect, meaning that *the critical speed* - it is the lowest speed of the vehicle, where hunting is possible - will decrease with increasing values of the 'effective conicity' for fixed values of all other parameters.

The wheel profile therefore has a strong influence on the dynamics of the vehicle. With wheel wear, the 'effective conicity' increases. Attempts have been made to design wheel profiles with an 'effective conicity' that is independent of - or very little dependent on the wear. These wheel profiles must have a profile that resembles the rail profile in order to distribute the wear more evenly. It means that the new profile is made to look like a worn one or - in other words - in order to prevent a destabilizing effect of the wear the stability properties are aggravated to begin with!

In the ideal case of rigid wheels and rails a wheel and a rail will have contact in one or more points. 'Wheel lift off' occasionally happens in reality, but it should be avoided since it can lead to a derailment. 'Wheel lift off' is an exception that is not considered in these notes. Since the effective conicity has a big influence on the dynamics of the vehicle its instantaneous value must be known. The values change with the lateral and yaw position of the wheel relative to the rail. When the rail and wheel profiles are known, the position of a contact point can be calculated as a function of the lateral position and the yaw angle of the wheel set.

The geometric variables that depend on the lateral position  $y$  of the wheel relative to the rail are:

- $r_L$  - instantaneous rolling radius of the left wheel
- $r_R$  - instantaneous rolling radius of the right wheel
- $z_L$  - instantaneous height of the contact point on the left rail
- $z_R$  - instantaneous height of the contact point on the right rail
- $\delta_L$  - contact angle of the left wheel
- $\delta_R$  - contact angle of the right wheel
- $\phi_W$  - roll angle of the wheel set with respect to the plane of the rails.

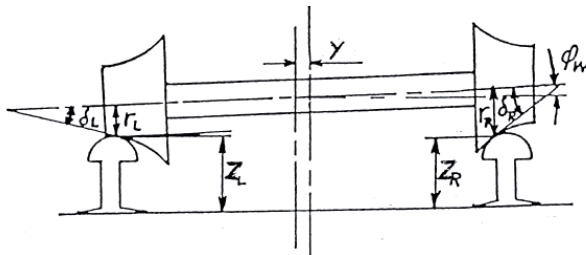


Figure 2. The wheel/rail parameters, rear view

These constrained variables are illustrated on Figure 2.

The variations of the position of the rails will of course influence the dynamics of the vehicle. Since the variations usually are small and the feed-back negligible the variations will enter the model equations as an external forcing.

## 1.2 Dynamics

In order to model the dynamics we shall adopt two different points of view. When we derive the relations between the contact forces and the contact geometry it is necessary to take the following aspects into account:

1. The contact between the wheel and the rail occurs over a finite surface that may not be plane. This is due to the compliance of the bodies in contact
2. The relative motion of different points in this surface may differ
3. The stress distributions across the surface are not uniform.

On the other hand we want to avoid to treat the two bodies as continua. It is therefore desirable to consider the contact as if:

1. the contact occurs at a single point
2. the wheel and the rail have a relative velocity at this point
3. this relative velocity lies in a plane (*the contact plane*) that is tangent to the bodies at the contact point as long as the two bodies remain in contact.

In general the contact point lies in a vertical plane that contains the centre line of the wheel set, and the contact plane is perpendicular to that plane. When the wheel set yaws, the contact point moves away from the vertical plane. The displacement is the so-called 'Vorverlagerung' (from German), that can be calculated by Heumann's method, Heumann (1950-1953). The dynamic effect of the 'Vorverlagerung' is important in curves, when the wheel set yaws. Experience has shown, however, that the 'Vorverlagerung' may be neglected, when the radius of the curve is larger than around 600 m with the wheel and rail profiles that are used today.

When the vehicle moves, there is rolling contact in the wheel/rail contact surface. Under the load both the wheel and the rail will deform. Hertz's theory of deformation has yielded satisfactory results for the calculation of the contact surface, the normal force and the mutual penetration of the wheel and the rail *as long as the curvature of the contact surface is negligible*.

**Normal Forces.** Hertz considered two convex elastic bodies that are pressed together by a force  $N$ . He assumed that the dimensions of the contact surface are small compared with the radii of curvature of the two bodies in a neighbourhood of the contact point. Hertz then showed that the contact surface would be elliptic with the major semiaxis  $a$  and the minor semiaxis  $b$ . The essence of Hertz's results is that:

- the ratio  $a/b$  only depends on the curvatures of the two bodies
- $a \propto N^{1/3}$ ,  $(a \cdot b) \propto N^{2/3}$  independent of the tangential stress distribution
- the mutual penetration  $q \propto N^{2/3}$ .

The contact angles of the right and left wheel respectively can only be assumed to be equal when the wheel set is centred in a straight track with zero yaw. In general they will be different, and the sum of the normal contact forces will then have a lateral resultant. This sum of the lateral components of the normal forces is often referred to as *the gravitational stiffness*. It is approximately equal to the axle load multiplied by half the difference of the contact angles plus the wheel set roll angle for small contact angles. At larger contact angles trigonometric functions of the angles enter the expressions for these terms thereby making the lateral resultant force a nonlinear function of the contact angles.

In most cases the gravitational stiffness force acts in a direction to accelerate the wheel set towards its centred position. If the gravitational stiffness is large owing to a large change of the contact angle difference with lateral displacement, significant restoring forces will result. Thus the gravitational stiffness in general has a stabilizing influence.

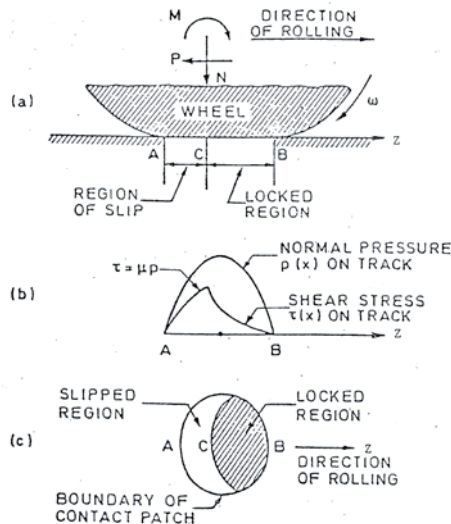
When one of the bodies is concave in a neighbourhood of the contact point, then the contact surface is not a plane surface and Hertz's theory must be used with caution. It may be used as a first approximation to determine qualitative features of the motion of the vehicle and its components. The calculated stresses in the contact surface will, however, most likely be erroneous. Recent work by Piotrowski and Kik (1998) takes the curvature of the contact surface into account in the stress calculations. Hoffmann (2006) made in his thesis an interesting comparison between the calculations of the dynamics of a two-axle freight wagon with application of first Hertz's theory and second Kik's RSGEO routine that takes the 'non-ellipticity' of the contact surface into account. He found that the results were qualitatively very similar, but the displacements using Kik's routine RSGEO were up to three times larger than the displacements found using Hertz's theory. The 'Kik results' are according to earlier experience very close to the real values.

If two contact points exist simultaneously in a wheel/rail combination, the calculation of the contact stresses becomes more difficult. Sauvage and Pascal (1990) and Pascal and Sauvage (1992) have formulated approximative methods to handle the problem, whereby the problem is reduced to the determination of the stresses in one fictive contact point. The contact point may also jump from one position to another under the influence of a lateral displacement of the wheel set. The jump introduces a discontinuity in the dynamic system, and we shall treat the numerical handling of such non-smooth problems in section 7.2. If, however, the compliance of the wheel and the rail is taken into account, then the growth of the second contact surface and the decay of the first are continuous functions of the displacement. The discontinuous rigid body problem is then reduced to the aforementioned two-point problem.

Several numerical routines have been made that take the 'non-ellipticity' and two-point contact into account. Kalker's program CONTACT has won international approval as the most accurate routine for the determination of the contact surfaces and calculation of the contact forces. Its disadvantage is the long CPU-time needed to run the program. Other programs that may be less accurate but need less CPU-time are commercially available. It is highly recommended to use these routines or CONTACT for vehicle dynamic investigations in future. I shall not give any specific recommendations, but only refer to the existing literature and the software providers.

In a survey lecture Piotrowski and Chollet (2005) described the approximative solution methods as well as later work in the area. The most recent contribution to the handling of the contact problem is by Quost et al. (2006).

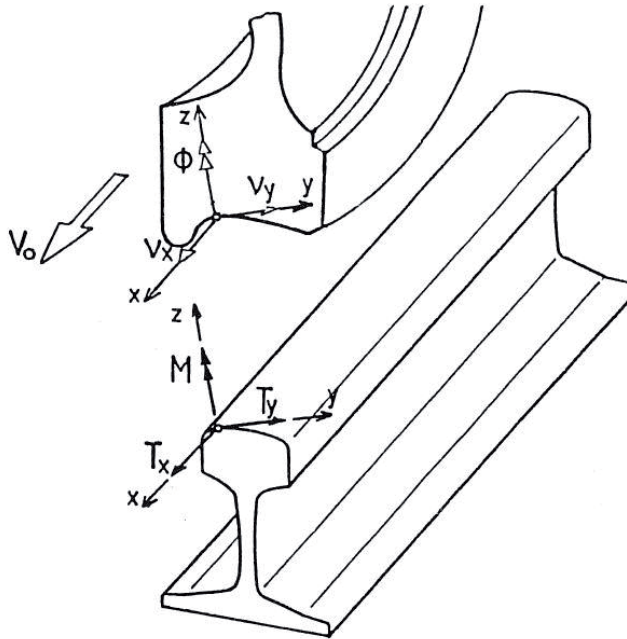
**Tangential Forces.** In rolling - not sliding - contact the contact surface is divided into two parts that differ in the kind of contact (See Figure 3). In the first part the two bodies stick to each other and the tangential deformations are purely elastic. In the second part the two bodies slide against each other, and Coulomb's friction law applied in this part yields satisfactory results for the tangential stresses in that part of the contact surface. The elastic deformation and the sliding together result in a finite sliding velocity in the contact surface. It is denoted *the creep*. The creep normalized with the speed of the vehicle is denoted *the creepage*. The effect of the creep is that the resulting relative speed of the wheel in the contact surface always is different from zero, when the vehicle moves. That of course is the source of the wear on both the rail and the wheel treads and flanges.



**Figure 3.** The wheel/rail contact surface

While the normal load is very important for the long-term dynamic deterioration of the track, it is the tangential forces that are most important for the vehicle dynamics. The creep creates tangential forces that are denoted *the creep forces*. The creep forces have two components in the contact surface: One is longitudinal in the direction of the motion of the wheel set, and the other is lateral. It is not necessarily horizontal due to the contact angle. Perpendicular to the contact surface is a third component: The spin creep  $\Phi$  which exerts a torque  $M$  (see figure 4). The torque is created by the resultant of the creep forces, which has a point of attack that in general differs from the centre of the contact ellipse and depends on the contact angle. When the contact angle is zero, then the spin creep is zero. When the contact angles of the left and right wheel of the wheel

set are equal, then the spin creeps of the wheels are equal and opposite. The resulting spin creep of the wheel set is then zero.



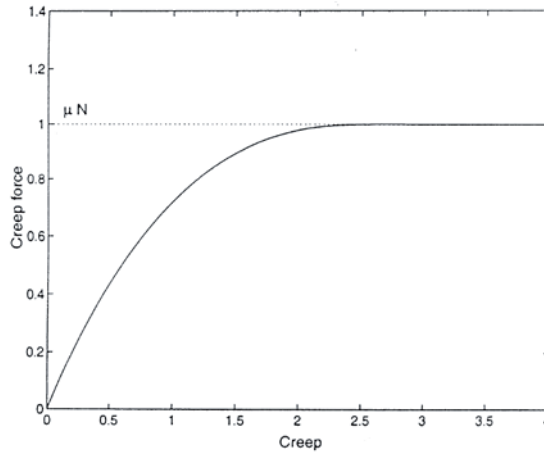
**Figure 4.** The wheel/rail creepages and creep forces

The problem of calculating the creepage and the creep forces is highly nonlinear, and it has been treated by several scientists. The main reference to this problem is Kalker. In a series of publications starting with his Ph.D. thesis Kalker has formulated and solved the rolling contact problem, and he has made computer programs of increasing accuracy and CPU time for the numerical computation of the stresses, the resulting contact forces and the spin torque. Kalker's most accurate program is the earlier mentioned program, CONTACT. It is generally accepted as the most accurate program for the calculation of the wheel/rail contact forces and torque. A formulation of the classical dynamic contact problem has been made by Chudzikiewicz (1991). A formulation of an extension of the classical theory to Non-Steady-State and Non-Hertzian problems is given by Knothe et al. (1991).

Other models of the contact forces versus the creepage have been formulated. The most wide spread one is the method by Shen, Hedrick and Elkins, since it is easy to implement. It yields good results as long as the spin creep is small. Shen et al. (1984) present a survey of these models.

A more recent survey of the models has been written by Kalker (1991). A recent contribution by Polach (2000) that is not included in Kalker (1991) describes a fast and for most applications sufficiently accurate computer program for the calculation of the





**Figure 5.** The general form of the dependence of the normalized creep force,  $F$ , on the creepage,  $c$

wheel/rail contact forces.

The general shape of the creep force normalized by the Coulomb friction force,  $F$ , versus the creepage,  $c$ , is shown on figure 5. It is clearly nonlinear and has the form of a characteristic of a weakening damper. The curve approaches the Coulomb friction limit for large creepage, which corresponds to pure sliding or wheel slip. For still larger creepages the curve drops off again. This is the cause of a hysteresis effect that is well-known from wheel sets under high tractive torques. Under traction so-called 'macro-slip' oscillations develop and under braking the vehicle braking force decreases, when the wheels are blocked.

The lateral accelerations of the vehicles are the most annoying disturbances. In order to calculate the lateral displacements and accelerations we must of course know the lateral creep forces. It must be remembered, however, that the longitudinal creepage has a reciprocal effect on the lateral creep forces. Furthermore the spin creepage produces a lateral force. This contribution to the total lateral force reduces the longitudinal part of the total creep force that is needed for the traction - sometimes quite substantially. This is the cause of the wheel slip of driven axles at high speed that has been observed in real life. The axle torque at high speed of the vehicle is smaller than under the initial acceleration, where high longitudinal creep forces are produced, so the 'high speed wheel slip' has puzzled the engineers for decades. Altogether there exists a rather complicated nonlinear relationship between all the creepages and the components of the creep forces.

Although Kalker's theory is called 'exact', it is important to be aware of Kalker's assumptions. Birkedal Nielsen (1998) discusses Kalker's assumptions and extends the theory, but Birkedal's own results are only two-dimensional. Birkedal investigates the effect of a speed dependent adhesion coefficient. His polynomial model of the contact surfaces allows very small characteristic wavelengths of the track irregularities such as the dreaded small wavelength corrugations. His model also can be used to investigate *non-*

*steady* two-dimensional rolling contact problems. The extension to three-dimensional contact problems *without spin creep* is conceptionally simple. It is done by extending the surface integral in Birkedal's theory over the entire contact ellipse. The application of the theory to the general three-dimensional contact problems is, however, still an open problem.

Recent contributions treat the wheel and the rail as elastic continua and solve the theoretical contact problem numerically using finite elements. Thereby an improved correspondance between the theoretical and measured results as well as a deeper understanding of the process have been achieved. The costs measured in CPU-time, however, are so high that the programs are less useful for most vehicle dynamics investigations. The interested reader is referred to Damme et al. (2003) and the references therein.

The wheel/rail contact geometry and the wheel/rail contact forces constitute the basis for all railway vehicle dynamic investigations including wear and fatigue. The problems to be solved are so complicated that a numerical analysis is mandatory for even the approximate solutions. The methods that use finite element methods and Kalker's program, CONTACT, consume so much CPU time that they hardly can be used implicitly as subroutines in any railway dynamic programs. Several applications of CONTACT exist, however, in which the possible rail/wheel contact configurations and forces versus the independent variables are calculated and tabulated a priori. The positions and the forces that are needed in the dynamic system for the calculation of the next time step are then found by interpolations of the tabulated values and used as external inputs in the dynamic system. It is recommended to use either splines or similar polynomial forms for the interpolation, since the non-smooth linear interpolation may create numerical problems. A higher accuracy can be achieved through the use of implicit interpolation schemes.

The approximations by Shen et al. (1984) and Polach (2000) are easy to implement and work so fast that they easily may be implemented implicitly in a numerical program for the determination of the dynamics of a railway vehicle. The interested reader may find the expressions to be programmed or even the routine itself in the references Shen et al. (1984) and Polach (2000).

## 2 Nonlinear Railway Vehicle Dynamics

### 2.1 Introduction

The vehicle dynamic problems are parameter dependent. Examples of parameters are the adhesion coefficient in the rail/wheel contact surface, spring stiffnesses in the suspensions and the speed of the vehicle. The dynamics of parameter dependent systems depend on the values of the parameters. If the parameters vary during operation, then the dynamics can change qualitatively at certain *critical* parameter values. In railway vehicle dynamics the speed,  $V$ , of the vehicle is an externally controlled parameter that vary during the operation. The investigation of the dynamics of a vehicle therefore must start with a search for the possible steady states of the dynamic system and the critical speeds where the dynamics may change qualitatively.

All railway vehicles known to the author have a *critical speed*. They may have more,

but it is only the lowest, which is called *critical*. At speeds below the critical speed the irregular motions of the vehicle will exclusively be responses to the track geometry and its irregularities. We neglect aerodynamic forces in these notes. Above the critical speed the vehicle oscillates in the plane of the track. Through coupling in the suspensions the oscillations may spread to roll and pitch motions. This is the well-known *hunting*. The critical speed depends on many parameters and its determination is not trivial. The railway industry tries to design the vehicles in such a way that the critical speed lies well above the maximum operational speed of the vehicle. The questions of initial costs, maintenance costs and robustness do, however, play an important role. The suspensions that guarantee a high critical speed are very sophisticated - nowadays active - systems that do not satisfy the cost criteria for every type of vehicle. Therefore the cheap designs for freight wagons most often have an operational speed range that includes the critical speed.

It has been assumed for more than a century that the determination of a critical parameter value in dynamic problems can be formulated mathematically as a stability problem. This is not true in general and for railway vehicles as a group of dynamic systems in particular. That is connected with the nonlinear nature of vehicle dynamics, and we must therefore take a look at the basics of nonlinear dynamics in the following chapters.

## 2.2 On nonlinear dynamic problems in general

In general all kinematical and constitutive relations in vehicle system dynamics are nonlinear. The modelling of such systems therefore leads to nonlinear dynamic systems, which must be investigated with respect to their properties and evolution.

It is characteristic for vehicle dynamic systems that they contain elements, which dissipate energy. Therefore we consider *only dissipative and non-conservative systems* in these notes.

The traditional way of solving a nonlinear, dissipative dynamic problem is to solve 'the corresponding' linear dynamic problem first and then apply perturbation theory by making an expansion of the solution of the fully nonlinear problem around the solution to the linear problem. This method has often led to satisfactory results, although the solution is approximate and the error will grow with the expansion parameter or coordinate. It is tacitly assumed that the solution can be approximated to the first order by the unique solution of a linear problem and the nonlinear influence can be found by addition of a correction - the terms in the perturbation expansion - to the solution of the linear problem. The first assumption is principally wrong! A nonlinear dynamic problem may have several different steady state solutions - even though it is dissipative. (See an example of multiple steady state solutions on Figure 6). In contrast the solution of a linear dissipative system without forcing will always tend to zero. With forcing it tends to the particular integral defined by the forcing independently of the initial conditions. Therefore the solution to the linear problem may be wrong as a first approximation to the nonlinear problem. The other assumption is, that the solution to the nonlinear problem can be approximated by a sum of two parts: The solution to the linear problem and a series expansion. That method will lead to a good approximation of a solution to

the nonlinear problem in a certain parameter interval provided the solution of the linear problem is a good first approximation - i.e. the first assumption holds - and the 'correct' perturbation method is used.

In vehicle system dynamics the dynamic problem is most often first to find the steady state(s), next to determine their stability and third to determine the behaviour of the vehicle when changes between the steady states, e.g. by braking a vehicle, occur. Therefore the fundamental problem is first to find the stationary and other steady states and their stability. *The analysis of a nonlinear dynamic problem must therefore always begin with a search for all the steady states depending on the parameters in the problem by solution of the full nonlinear problem.* In an appropriate frame at least one of the steady states can often be found easily, but that is only the beginning.

In the simplest nonlinear problems some steady state solutions can be found analytically but most often numerical methods must be used. The use of numerical methods will be discussed in section 7. If multiple steady states exist, then they fortunately very often develop through so called *bifurcations* or branching off from other steady states, when a control parameter in the problem is varied. There are several types of bifurcations and the ones most relevant for vehicle dynamic systems are described in chapter 4.

This fact offers us a possibility to find multiple steady states by tracing already found steady state solutions by varying the control parameter through the values where bifurcations take place and then change to follow the new solution from the bifurcation in dependence on the control parameter. The fundamental steps in the solution of a nonlinear dynamic problem therefore consist of finding at least one steady state solution in a sufficiently large parameter interval and determine its stability in dependence on the parameter. Then find bifurcation points where different steady solutions will be found. The new solutions and their stability must then be determined in dependence on the control parameter. The new solutions may create new bifurcation points, which must be found, and the game starts all over in these new bifurcation points! Unfortunately there is no guarantee that *all* the steady state solutions will be determined in this way, but experience shows that the method is very efficient. The method is described in detail in section 7.3. Good engineering know-how is still needed to judge whether all possible states have been found or it is necessary to use other strategies to find possible missing states.

Nonlinear mechanics deals with the formulation or modelling of processes, while nonlinear dynamics is a mathematical discipline. The topics therefore are different. We shall deal with *nonlinear and deterministic dynamics* in these notes. Stochastic variables are excluded. In nonlinear problems *the existence* of solutions is the primary object of analysis in contrast to *stability*, which comes second. When several stable steady state solutions of a dissipative problem coexist, each and everyone will have a set of its own initial conditions with the property that all solutions determined by initial conditions from that set will tend to that particular stable solution. Such a set of initial conditions is called a *basin of attraction*. This fact has an important implication. If the dynamic problem has multiple steady states and the corresponding physical vehicle in the real world is subject to uncontrolled initial conditions, then finite probabilities exist for the vehicle to tend to any one of the stable steady states. In other words: When several stable steady state solutions of a dissipative dynamic problem coexist then the vehicle is

*unstable to finite disturbances.*

The mathematical properties of nonlinear systems are fundamentally different from the mathematical properties of linear systems, and the nonlinear systems must be treated accordingly! An important difference between the properties of linear and nonlinear systems is that the principle of superposition does not apply to nonlinear systems. The principle of superposition is the basis of integral transform solution methods. Another difference is the effect of the initial conditions on the steady state solution of an initial value problem. A linear, dissipative problem has only one steady state solution independent of the initial conditions, but a nonlinear problem may have more steady state solutions, where each one depends uniquely on the applied initial conditions. The existence of more steady state solutions of a nonlinear problem is demonstrated in the following easy example of a nonlinear initial value problem. (From Thompson and Stewart (2002, p.23)).

Consider the nonlinear differential equation:

$$2.56(d^2x/dt^2) + 0.32(dx/dt) + x + 0.05x^3 = 2.5\cos t; \quad (2.1)$$

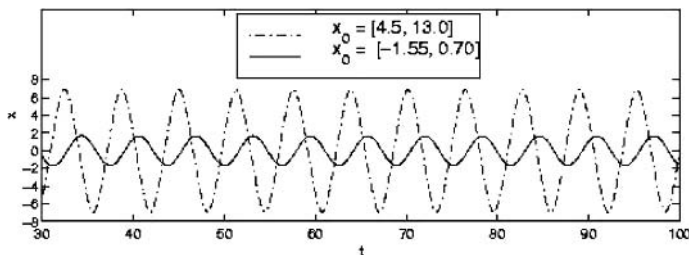
with the initial conditions:

$$x(0) = 4.50, \quad dx/dt(0) = 13.00; \quad (2.2)$$

and

$$x(0) = -1.55, \quad dx/dt(0) = 0.70; \quad (2.3)$$

The two different initial value problems (2.1)+(2.2) and (2.1)+(2.3) lead to two different steady solutions - called *attractors*. They are found numerically using the Runge-Kutta solver rk45 from MAPLE 9.5. The attractors are depicted on figure 6.



**Figure 6.** The two solutions of (2.1)+(2.2) and (2.1)+(2.3)

It is easily seen that although the excitation is a periodic function with the amplitude 2.50 and the system is damped, then there exists a steady state large amplitude solution to the differential equation in addition to the expected steady state solution with amplitude less than two. This is an example of *coexisting steady state solutions*. When the

corresponding linear problem is solved without the  $x^3$ -term, only a low amplitude steady state periodic solution exists for both sets of initial conditions.

Furthermore oscillation frequencies are amplitude dependent in nonlinear dynamics. As a consequence hereof *the concept of resonance such as it is used in linear theory must be reconsidered and revised.*

In order to simplify the nonlinear dynamic problem, the system may be idealized within certain limits. The simplest assumption is that all the parameters are independent of the state of the system, but it is wrong for all real mechanical systems. In many cases, however, it is possible to find intervals of the action in which certain parameters in practice (i.e. within a given small tolerance) remain constant. It may for example happen that the action of a damper in the problem of interest is limited to such a small velocity interval that its nonlinear character may be neglected. Such simplifications in the theoretical model must, however, be made with great caution, because they may change the solutions of the problem *qualitatively and/or quantitatively* and important properties of the solution may be lost.

In the following we shall introduce the basic notions of a *state space* and *trajectories*, because they lead to a visualization of the solutions of dynamic problems, their uniqueness (or lack of same) and stability. We shall also present the conditions under which a linearization of a nonlinear operator (problem) is permitted and indicate how the solution of the linearized problem can give valuable information about the local properties of the nonlinear problem. We shall define and examine stationary and other steady state solutions, their stability and bifurcations. Finally properties of solutions that cannot exist in linear systems - such as chaos - will be described and their effects briefly discussed.

### 2.3 The Formulation of a Railway Vehicle Dynamic Problem

The problems in theoretical vehicle dynamics are formulated as nonlinear dynamic multi-body problems. The rolling wheel/rail contact as well as the more rarely applied sliding contact are the most important features of a vehicle dynamic system, and the constitutive laws for these contacts are nonlinear. In a multi-body system the single bodies are interconnected, and several of the connections - e.g. springs and dampers - obey nonlinear constitutive laws. Guides introduce constraints, which often are formulated by algebraic or transcendental equations, so the problems in vehicle dynamics must often be formulated as a system of differential-algebraic equations (DAE). Bumper stops and dry friction contact are important nonlinearities that introduce discontinuities. The discontinuities violate the assumptions of the mathematical theorems, which form the basis of many methods that are applied for the solution of systems of differential-algebraic equations.

In vehicle dynamics the analysis of the problem is most often formulated as an initial value problem for a system of ordinary differential equations with the time  $t$  as the single independent variable. The system of differential equations may be combined with algebraic or transcendental equations that express certain kinematical constraints e.g. the geometric contact between a wheel set and the rails. The algebraic equations are in general nonlinear, and they must be solved together with the system of differential equations as one system of differential-algebraic equations. Boundary value problems

for partial differential equations or equations with time lag may, however, also enter the dynamic system, but they will not be considered in these notes.

### 3 On the Geometry of Nonlinear Autonomous Dynamic Systems

#### 3.1 Some Definitions

The geometric theory of nonlinear dynamic systems provides the best access to an understanding of the behaviour of such systems. A dynamic system is a system of first order ordinary differential equations that often is connected with a system of algebraic or transcendental equations, which define constraints in the system. *The state of the system* is defined as the set of positions and velocities of each and every element of the vehicle in a suitable coordinate system. Both the positions and the velocities are treated as dependent variables in the dynamic system with time  $t$  as the independent variable. The dimension  $N$  of the state space equals the number of dependent variables of the dynamic system. The mathematical problem consists of finding the state of the system under external action as a function of the time and in dependence on the parameters. In the following section we shall assume that the parameters are all constant.

The dependent variables  $x_i, i = 1, 2, \dots, N$  are called *state variables* and they are all real.  $\mathbf{x}$  is the  $N$ -dimensional vector  $x_i, i = 1, 2, \dots, N$ . The time,  $t$ , is the independent variable. If the time  $t$  does not enter the dynamic system explicitly, the system is called *autonomous*. If the time  $t$  enters the dynamic system explicitly, or the parameters depend on time, then the dynamic system will be *non-autonomous*. An  $N$ -dimensional non-autonomous dynamic system can be converted into an  $(N+1)$ -dimensional autonomous system by the introduction of the time  $t$  as an additional dependent variable by the phase equation  $dt/d\tau = 1$ , and consider  $\tau$  the new independent time variable. We shall therefore only consider autonomous systems in the following. If the parameters depend on the state of the dynamic system, then the parameters will give rise to additional nonlinearities.

We denote  $\frac{dx_i}{dt}$  by  $\dot{x}_i$  and correspondingly  $\frac{d\mathbf{x}}{dt}$  by  $\dot{\mathbf{x}}$ .

$\lambda = \lambda_j, j=1,2,\dots,M$  denotes a set of  $M$  real valued parameters. A *parameter dependent nonlinear dynamic system* can then be written in the form:

$$\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}; \lambda). \quad (3.1)$$

$\mathbf{F}$  is an  $N$ -dimensional real valued ‘sufficiently smooth’ nonlinear vector function of the  $N$  dependent variables  $x_i$ . By *function* we always mean an expression from which the value is determined *uniquely* by the arguments. If uniqueness is not guaranteed we speak of a *relation*. An example of a relation is the expression  $f = \sqrt{x}$ , since there exist two real values of  $f$  for every positive value of  $x$ . When we deal with functions we call the set of all permissible values of the arguments  $x_i$  *the domain of the function*  $\mathbf{F}$ . The set of values that the function can assume for all permissible variations of the arguments is called *the range of the function*  $\mathbf{F}$ .

A mechanical system with  $K$  degrees of freedom ( $K$  DOF) will normally be modelled mathematically by a system of  $K$  ordinary differential equations of second order in the dependent variables  $q_i$ . The trivial substitution  $q_i = x_{2i-1}$  and  $\dot{q}_i = x_{2i}$  will however

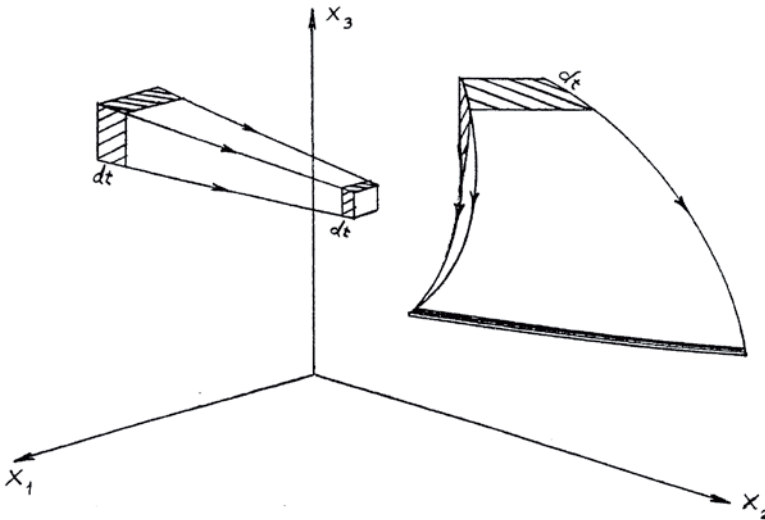
transform the system of  $K$  second order differential equations into a system of  $2K$  first order differential equations and we then set  $N$  equal to  $2K$ .

We now introduce a cartesian coordinate system with the state variables  $x_i(t)$  as the coordinates. It will be  $N$ -dimensional and it is denoted *the state space* or *phase space*. We can plot the values  $x_i(t)$  that satisfy eqn.(3.1) for fixed parameter values  $\lambda_j$  in the phase space. We then obtain a set of curves, depending on the initial conditions, which are called *trajectories* or *integral curves* corresponding to eqn.(3.1).  $t$  is a curve parameter and associated with each trajectory is a direction showing how the state of the system changes as time increases. The complete figure is called *the phase diagram* or *phase portrait* for the system. A point in the phase space is called an *image point* and the trajectory through an image point is often termed a *phase trajectory* in order to distinguish it from the trajectory of the motions in the physical world. The speed of an image point is called *the phase speed* not to be confused with the physical speed of the elements in the system.

In autonomous systems *trajectories can not cross each other*, meaning that the motion of the system is determined *uniquely* by any point on a trajectory. An *equilibrium point* is an image point that does not change with time, so the phase trajectory is just a point. It corresponds to a *steady state* of the dynamic system, and it is a *stationary solution*. A point for which the slope of the trajectory is not defined is called a *singular point*, all other points are called *regular points*. An equilibrium point is a singular point.

Since a periodic solution will always return to any given state with its period  $T$ , *the trajectory of a periodic solution must be a closed curve*.

Dissipative autonomous systems are *contracting in the phase space* in the sense that a small volume bounded by a bundle of trajectories and lying between  $t_0$  and  $t_0 + dt$  will



**Figure 7.** Examples of trajectories of a dissipative dynamic system in state space



decrease when it moves along its bundle of trajectories with growing  $t$  (see figure 7 for two examples).

### 3.2 Stability Concepts

We now introduce the four most important concepts of stability. First we define the stability of equilibrium points or stationary solutions. Given

$$\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}; \boldsymbol{\lambda}) \quad (3.2)$$

with the equilibrium point  $\mathbf{0}$ . It means that

$$\mathbf{F}(\mathbf{0}; \boldsymbol{\lambda}) = \mathbf{0}. \quad (3.3)$$

A simple coordinate transformation will take any one point in the state space to  $\mathbf{0}$ . For equilibrium points we first define *Local Stability* or *Lyapunov Stability*.

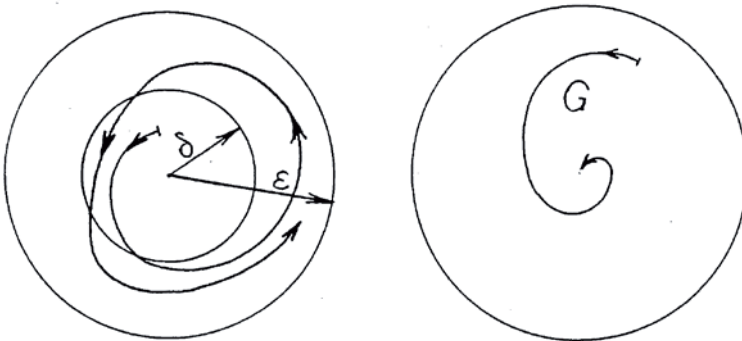
**Definition 1.** *If there for all  $\epsilon > 0$  exist  $\delta > 0$  (to every  $\delta$  there corresponds an  $\epsilon$ ) such that  $\|\mathbf{a}\| < \delta \rightarrow \|\mathbf{x}(t)\| < \epsilon$  for all  $t$ , where  $\mathbf{a}$  denotes an initial condition and  $\|\cdot\|$  denotes a suitable norm (or measure) then the point  $\mathbf{0}$  is locally stable or Lyapunov stable.*

The definition states that a trajectory, which starts inside a ball with radius  $\delta$  and centre in  $\mathbf{0}$  (the domain needs not be spherical) never will leave the ball with radius  $\epsilon$  and centre in  $\mathbf{0}$  (the domain is not necessarily spherical) (see Figure 8).

**Definition 2.** *If  $\mathbf{0}$  is locally stable for all trajectories that start inside a domain  $G$ , which contains  $\mathbf{0}$ , and the following relation in addition holds*

$$\|\mathbf{x}(t)\| \rightarrow \mathbf{0} \text{ for } t \rightarrow \infty$$

*then  $\mathbf{0}$  is asymptotically stable.*



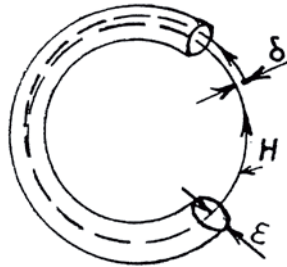
**Figure 8.** Trajectories near the stable point  $\mathbf{0}$ . Left: Locally stable and right: Asymptotically stable

It is the second and most used concept of stability for dissipative systems. The definition states that  $\mathbf{0}$  is asymptotically stable if all trajectories that start in  $G$  or enter  $G$  ultimately will come arbitrarily close to  $\mathbf{0}$ .

Next we define *Orbital Stability*. We now consider an orbit or closed trajectory  $H$  and a tube of radius  $\epsilon$  around  $H$ . Please remember that a closed trajectory in phase space represents a periodic solution in the physical space.

**Definition 3.** *If all paths starting from initial points  $\mathbf{a}^*$  within a distance  $\delta$  from any point  $\mathbf{a}$  on  $H$  remain permanently within a tube around  $H$  of radius  $\epsilon$  for  $t \rightarrow \infty$  then the orbit  $H$  is orbitally stable.*

If all the paths come arbitrarily close to  $H$  for  $t \rightarrow \infty$  then the orbit  $H$  is *asymptotically orbitally stable* (see Figure 9). This definition can be extended to other steady state solutions of the dynamic system.



**Figure 9.** An asymptotically orbitally stable trajectory

Points or paths that are not stable (locally or orbitally) are *unstable*. They are unstable under *infinitesimally small disturbances*.

All the stability concepts mentioned above refer to stability with respect to disturbances of *initial conditions*. The last concept of stability we shall use, refer to disturbances of the vector function  $\mathbf{F}(\mathbf{x}; \boldsymbol{\lambda})$ .

We therefore finally define *structural stability*.

**Definition 4.** *A dynamic system is structurally stable if small disturbances of the vector function  $\mathbf{F}(\mathbf{x}; \boldsymbol{\lambda})$  lead to equivalent phase diagrams.*

An observable phenomenon should always be modelled by a structurally stable system, because small errors in the modelling and fluctuations in the physical system are unavoidable. A conservative dynamic system is structurally unstable according to this definition, since the addition of an ever so small damping will change the phase diagram qualitatively.

For the investigations of asymptotic stability of equilibrium points *Lyapunov's first method* may be applied. With no loss of generality we shall again assume that the equilibrium point is  $\mathbf{0}$ . Given the autonomous system:

$$\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}; \boldsymbol{\lambda}) \quad (3.4)$$

with  $\mathbf{F}(\mathbf{0}; \boldsymbol{\lambda}) = \mathbf{0}$ .

We assume that the Jacobian

$$\mathbf{J} = [\partial f_i / \partial x_j], \quad i, j = 1, 2, \dots, N$$

exists and provides the matrix  $\mathbf{J}_0$  in  $\mathbf{x} = \mathbf{0}$ . We then write eqn.(3.4) in the form:

$$\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}; \boldsymbol{\lambda}) = \mathbf{J}_0 \mathbf{x} + \mathbf{G}(\mathbf{x}; \boldsymbol{\lambda}), \quad \mathbf{J}_0 = \mathbf{J}_0(\boldsymbol{\lambda}).$$

The equation

$$\dot{\mathbf{x}} = \mathbf{J}_0 \mathbf{x} \quad (3.5)$$

is the eqn.(3.4) linearized around  $\mathbf{0}$ , and  $\mathbf{G}(\mathbf{x}; \boldsymbol{\lambda})$  is a remainder, which contains all the nonlinear contributions.

**Theorem 1.** Let  $\mathbf{0}$  be an equilibrium point of  $\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}; \boldsymbol{\lambda})$ .

1. If the Jacobian  $\mathbf{J}_0$  has neither zero nor purely imaginary eigenvalues and
2.  $\|\mathbf{G}(\mathbf{x}; \boldsymbol{\lambda})\| / \|\mathbf{x}\| \rightarrow 0$  for  $\|\mathbf{x}\| \rightarrow 0$ ,

then it is a necessary and sufficient condition for asymptotic stability of  $\mathbf{0}$  that all solutions of  $\dot{\mathbf{x}} = \mathbf{J}_0 \mathbf{x}$  are locally stable.

In other words: Under the conditions stated *the linearized system* determines the stability of the full nonlinear system in a certain neighbourhood of  $\mathbf{0}$ . Vehicle dynamic systems are dissipative systems, and local stability of  $\dot{\mathbf{x}} = \mathbf{J}_0 \mathbf{x}$  is therefore guaranteed, when all the eigenvalues of  $\mathbf{J}_0$  have negative real parts. In definition 1,  $\delta$  may be very small, and in theorem 1 the condition 2 may only be satisfied in a very small neighbourhood of  $\mathbf{0}$ . Therefore the stability is local in the sense that the stability of the trivial solution,  $\mathbf{x} = \mathbf{0}$ , is only guaranteed for infinitesimally small disturbances. The same limitation holds for linear stability analyses.

In order to linearize a dynamic system it must be determined by a function in order to guarantee uniqueness of the point you linearize around. When the function is expanded in a series around the point of expansion, the linear part must be the lowest order contribution when the variables approach the point of expansion. Therefore the coefficients of the linear terms in the expansion must be finite and non-zero. Therefore  $\mathbf{F}(\mathbf{x}; \boldsymbol{\lambda})$  in (3.1) cannot be linearized around 0 for instance when  $\mathbf{F}(\mathbf{x}) = \sqrt{x}$ ,  $x^2$  or  $\cos x$ .

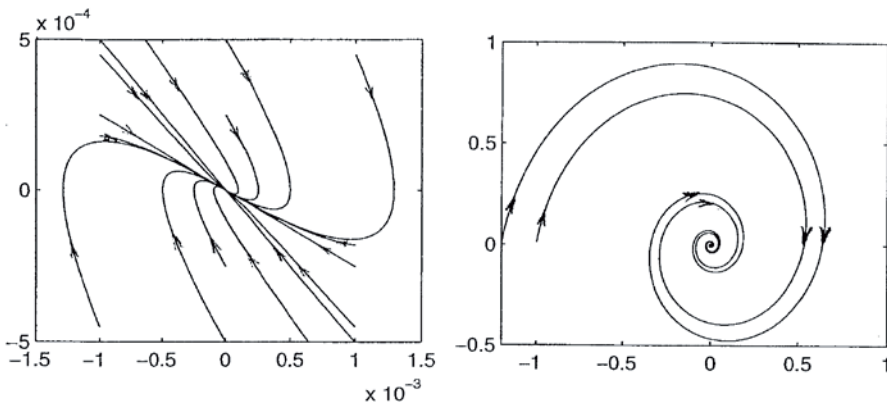
### 3.3 Equilibrium solutions

We now consider equilibrium points and cycles of dissipative, autonomous problems with the parameters fixed.

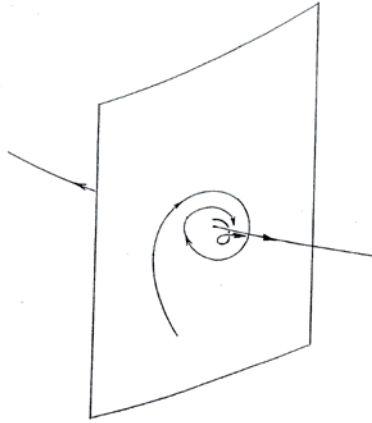
$$\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}). \quad (3.6)$$

**Equilibrium points.** We shall only consider equilibrium points for which theorem 1 holds. The equilibrium points and their stability are then characterized by the sign of the eigenvalues of the Jacobian if the eigenvalues are real, and by the sign of the real part of the eigenvalues if the eigenvalues are complex. Since we have assumed that the state variables as well as the vector function  $\mathbf{F}(\mathbf{x})$  are real, then the complex eigenvalues will always appear as complex conjugate pairs. If all the eigenvalues are real and of the same sign then the equilibrium point is called a *nodal point* or a *node*. If all the eigenvalues are negative then the nodal point is asymptotically stable, otherwise it is unstable. An asymptotically stable nodal point for the simple harmonic oscillator corresponds to an over-damped motion. If some eigenvalues are complex then the equilibrium point is called a *spiral point* or a *spiral* or a *focus*. If the real parts of all the eigenvalues are negative then the spiral is asymptotically stable. An asymptotically stable spiral for the simple harmonic oscillator corresponds to an under-damped motion. Notice that the corresponding eigenvectors span directions - in the case of nodal behaviour - or planes in the case of spiral behaviour (see Figure 10). In state spaces of dimension three or higher an equilibrium point may have nodal properties along some directions and spiral properties in some surfaces.

If the real part of the eigenvalues have different signs then the equilibrium point is a *saddle point* or just a *saddle*. A saddle is *always unstable*, since the eigenvectors belonging to the eigenvalues with positive real parts are directions of repulsion (see Figure 11).



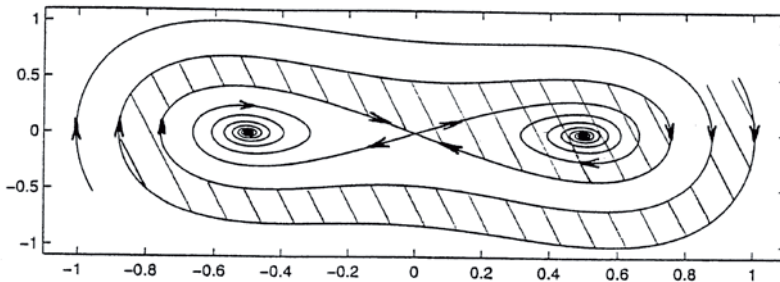
**Figure 10.** A stable node (left) and a stable spiral (right). If the directions of motion are reversed then the equilibrium points are unstable



**Figure 11.** An example of a saddle in a three-dimensional phase space. All saddles are always unstable

Notice that an equilibrium point is itself a trajectory, which consists of just that point. The trajectories that approach a stable equilibrium point will only get arbitrarily close to it in any finite time and never reach it! The theorem that trajectories in an autonomous system cannot cross each other thus still holds true.

The set of all the points on the trajectories that approach a stable equilibrium point is called *the basin of attraction* of that particular equilibrium point. When several stable equilibrium points coexist, it is particularly important to determine the basins of attraction for each of them, since one or more of the equilibrium points may be undesirable states (see Figure 12). Initial values that lead to the undesirable states should therefore be avoided.



**Figure 12.** Two stable spirals separated by a saddle. The "incoming" asymptotic trajectories (insets) split the domains of attraction of the two spirals

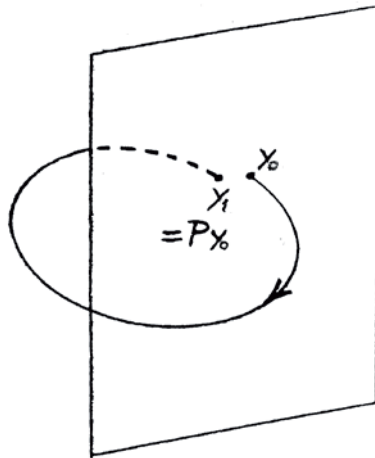
**Cycles.** Cycles are state space representations of periodic motions. They are closed curves in the state space because of the periodicity. Before we consider cycles in more detail it will be convenient to introduce the concept of a *discrete map* or just *map*. It is the discrete equivalent of our dynamic system eqn.(3.1).

An equation in the form:

$$\mathbf{y}_{n+1} = \mathbf{F}(\mathbf{y}_n; \boldsymbol{\lambda}), \quad n = 1, 2, \dots \tag{3.7}$$

maps the  $n$ th vector  $\mathbf{y}_n$  into the  $(n+1)$ st vector  $\mathbf{y}_{n+1}$  by the mapping function  $\mathbf{F}$ .  $\mathbf{F}$  is an  $N$ -dimensional real valued 'sufficiently smooth' nonlinear vector function of the  $N$  dependent variables  $y_n^i$ , ( $i = 1, 2, \dots, N$ ).

When we investigate the dynamics of periodic or other recurrent motions we shall use the *Poincaré map*. We introduce a plane in the state space and orient it in such a way that the trajectory intersects it transversally i.e. it is not a tangent plane to the trajectory at any point. Such a plane is called a *Poincaré section*. A cycle will intersect the Poincaré section at least twice but we now choose to consider only one of the intersection points  $\mathbf{y}_0$ . We now follow the trajectory from  $\mathbf{y}_0$  all the way around the trajectory until it intersects the Poincaré section again *in the same direction*. We call that point  $\mathbf{y}_1$ . It is the image of  $\mathbf{y}_0$  under the *Poincaré map*  $\mathbf{P}$ .  $\mathbf{P}$  thus takes the point  $\mathbf{y}_0$  into  $\mathbf{y}_1$ , and in the case of a cycle  $\mathbf{y}_1$  will in general be identical to  $\mathbf{y}_0$  so  $\mathbf{P}$  reduces to the identity map  $\mathbf{I}$ . The domain and range of the map  $\mathbf{P}$  lie in the Poincaré section only. You can think of the Poincaré map as a stroboscopic image of a point moving on a recurrent trajectory.



**Figure 13.** The Poincaré map

The Poincaré map transforms a periodic solution of equation (3.1) into a stationary solution in the Poincaré section of

$$\mathbf{y}_{n+1} = \mathbf{P}(\mathbf{y}_n; \boldsymbol{\lambda}), \quad n = 1, 2, \dots \tag{3.8}$$

An equilibrium or fixed point  $\mathbf{y}_0$  of the Poincaré map  $\mathbf{P}$  has the property:

$$\mathbf{y}_0 = \mathbf{P}^m(\mathbf{y}_0; \lambda), \quad m = 1, 2, \dots$$

where  $\mathbf{P}^m$  denotes the  $m$ th iterate of the Poincaré map  $\mathbf{P}$ . Without loss of generality we assume that  $\mathbf{y}_0 = \mathbf{0}$ . We can now apply the theory of stability of solutions of discrete maps to the equilibrium point for  $\mathbf{P}$ . In analogy with the theory for continuous systems a linearization  $\mathbf{P}_y$  of  $\mathbf{P}$  is calculated and evaluated at the equilibrium point  $\mathbf{0}$ . The corresponding linear map will then be

$$\mathbf{y}_{n+1} = \mathbf{C}\mathbf{y}_n, \quad (3.9)$$

where the elements of the matrix  $\mathbf{C}$  are the elements of  $\mathbf{P}_y$  evaluated at the equilibrium point. We now introduce the linear transformation

$$\mathbf{y} = \mathbf{A}\mathbf{z}$$

in equation (3.9) and obtain

$$\mathbf{A}\mathbf{z}_{n+1} = \mathbf{C}\mathbf{A}\mathbf{z}_n$$

We assume that  $\mathbf{A}$  is non-singular and multiply from the left with  $\mathbf{A}^{-1}$  and obtain

$$\mathbf{z}_{n+1} = \mathbf{M}\mathbf{z}_n, \quad \mathbf{M} = \mathbf{A}^{-1}\mathbf{C}\mathbf{A} \quad (3.10)$$

We choose  $\mathbf{A}$  so that  $\mathbf{M}$  has a Jordan canonical form. If the eigenvalues  $\rho_m$  of  $\mathbf{C}$  are distinct,  $\mathbf{M}$  is a diagonal matrix with entries  $\rho_1, \rho_2, \rho_3, \dots, \rho_N$ . Then the  $k$ th iterate of equation (3.10) can be rewritten as:

$$z_k^m = \rho_m^k z_0^m, \quad m = 1, 2, \dots, N \quad (3.11)$$

where  $z^m$  is the  $m$ th component of  $\mathbf{z}$ . It follows from equation (3.11) that for  $k \rightarrow \infty$ ,

$$z_k^m \rightarrow 0 \quad \text{if } |\rho_m| < 1$$

$$z_k^m \rightarrow \infty \quad \text{if } |\rho_m| > 1$$

$$z_k^m = z_0^m \quad \text{if } \rho_m = 1$$

$$z_{2k+1}^m = -z_0^m \quad \text{if } \rho_m = -1$$

$$z_{2k}^m = z_0^m \quad \text{if } \rho_m = -1.$$

We see that the location of the eigenvalues  $\rho_m$  of  $\mathbf{C}$  in the complex plane determines the stability of the equilibrium point  $\mathbf{y}_0$ . If all the eigenvalues of  $\mathbf{C}$  lie inside the unit circle in the complex plane the equilibrium point  $\mathbf{y}_0$  is asymptotically stable. If at least one of the eigenvalues lies outside the unit circle then  $\mathbf{y}_0$  is unstable. If at least one of the eigenvalues lies on the unit circle and all the rest inside, then a linear stability analysis will be inconclusive.

If  $0 < \rho_m < 1$  for all  $\rho_m$  then disturbances will decay monotonously, and if for at least one eigenvalue  $\rho_k$ ,  $-1 < \rho_k < 0$  then disturbances will decay in an oscillatory fashion. We call this behaviour *flipping*. If at least one eigenvalue is complex, then the disturbance will rotate around the equilibrium point while it decays or grows depending on whether all eigenvalues are inside the unit circle in the complex plane or not. A cycle is thus *orbitally asymptotically stable*, if all the eigenvalues of  $\mathbf{C}$  lie inside the unit circle in the complex plane.

As before, the set of all the points on the trajectories in the state space that approach a stable cycle is called *the basin of attraction* of that particular cycle. A stable steady state solution is called an *attractor*. It can be a point, a cycle or a higher dimensional or even fractal dimensional structure. An unstable stationary periodic solution is most often a *saddle-cycle*. When several attractors coexist, it is desirable to determine the basins of attraction for each of them in order to guard against unpleasant surprises caused by initial conditions that will lead to one of the undesirable attractors.

All that has been said above about the calculation of stability of cycles only serves to help the reader understand the concept and the relation between cycles in state space and Poincaré maps. It is not a recipe for a stability analysis of periodic solutions of nonlinear problems. The reason is that *the author has never seen any realistic nonlinear dynamic system*

$$\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}; \boldsymbol{\lambda})$$

with a periodic solution, where the corresponding Poincaré map

$$\mathbf{y}_{n+1} = \mathbf{P}(\mathbf{y}_n; \boldsymbol{\lambda}), \quad n = 1, 2, \dots$$

could be determined explicitly. Instead *Floquet theory* is often applied directly to the continuous problem equation (3.1). We shall however here present an easier method that is suitable for a numerical analysis.

Given a periodic solution to equation (3.1),  $\mathbf{x}_0(t; T_0)$  with the known period  $T_0$ . Choose a value of  $t$ , say  $t_0$  and calculate  $\mathbf{x}_0(t_0; T_0)$ . Put a plane S that contains  $\mathbf{x}_0(t_0; T_0)$  in the state space transversal to the trajectory of  $\mathbf{x}_0(t; T_0)$ . Add a small disturbance  $\mathbf{y}$  to  $\mathbf{x}_0(t_0; T_0)$  in S, so both  $\mathbf{x}_0(t_0; T_0)$  and  $\mathbf{x}_0(t_0; T_0) + \mathbf{y}$  lie in S. Calculate the distance  $d_0 = \|\mathbf{y}\|$  between  $\mathbf{x}_0(t_0; T_0)$  and  $\mathbf{x}_0(t_0; T_0) + \mathbf{y}$ . Integrate the dynamic problem equation (3.1) over one or a couple of periods  $T_0$  and obtain another point  $\mathbf{x}_1 = \mathbf{x}_0(t_0; T_0) + \mathbf{z}$  in H. Calculate the distance  $d_1 = \|\mathbf{z}\|$  between  $\mathbf{x}_0(t_0; T_0)$  and  $\mathbf{x}_0(t_0; T_0) + \mathbf{z}$ . Finally calculate the ratio  $\rho = d_1/d_0$ . If  $\rho < 1$  then the periodic solution  $\mathbf{x}_0(t; T_0)$  is *asymptotically orbitally stable*. Otherwise it is unstable.

In conclusion we shall note that recurrent solutions that are characterized by several periods also are possible in nonlinear dynamic systems. They are called *multiperiodic* solutions. In case all the periods are commensurate - i.e. the ratio between each pair of periods is a rational number - then the resulting motion is periodic. If that is not the case then the resulting motion is quasi-periodic.

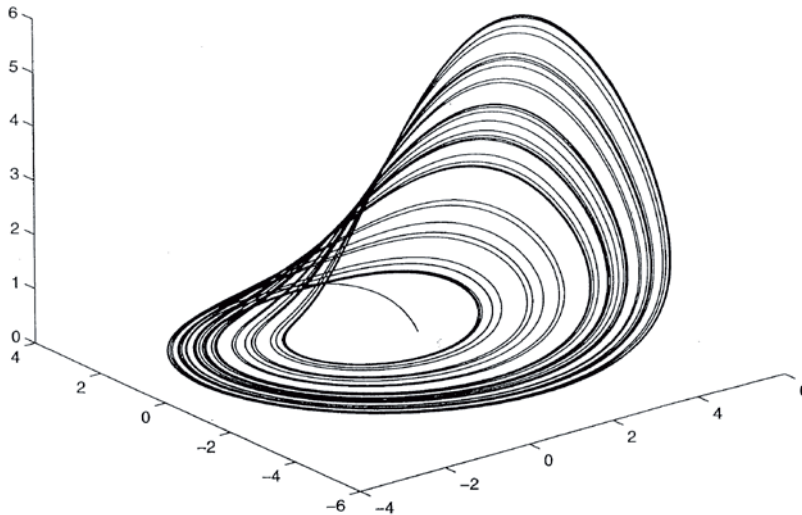
**Chaos.** It is a well known fact that deterministic nonlinear dynamic systems can behave erratically. We want here to make a clear distinction from random motion. We shall call



a process *random*, when uncertainties in the dynamic system appear as probabilistic relations or state variables. Such systems are not considered here.

The kind of erratic behaviour in deterministic systems is therefore of another character, and it is often called *chaos*. There exist both *chaotic attractors* - also called *strange attractors* - and *chaotic repellers*. Since only the attractors can be observed in real life situations we shall only describe those. There is a third important form of chaotic motion - *chaotic transients* - which will be discussed briefly later. Chaotic motion is *recurrent* but it is neither periodic nor quasi-periodic. Since it is recurrent but non-periodic the trajectories of a chaotic motion will 'return' to the neighbourhood of points they passed earlier without ever closing up at any point. Furthermore all trajectories on a chaotic attractor are orbitally unstable. They reject each other like saddle cycles, so in order to hold the bundle of all trajectories inside a bounded domain a *reinjection mechanism* must be at work in the dynamic process. That reinjection mechanism takes a bundle of trajectories and folds them back onto the other trajectories without touching any of them. This action is the cause of *mixing* in chaotic dynamics. In a three-dimensional phase space the attractor therefore consists of infinitely many infinitely thin 'layers' in which the trajectories are imbedded - a structure that can be described by the mathematical concept of a *fractal dimension*. The layers are separated and so close to each other that the attractor seems two-dimensional in the three-dimensional phase space.

In Figure 7 (right hand side) is shown an example of a set of trajectories that repel each other in a surface and contract in the direction normal to the surface. This action is typical for chaos. Due to the dissipation the volume elements in phase space must decrease with time - as earlier mentioned - but that property does not prohibit a



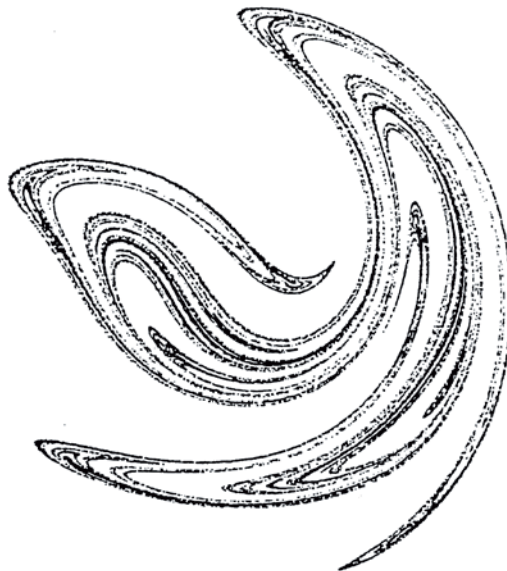
**Figure 14.** Rössler's chaotic attractor. It is an equilibrium solution of the dynamic system:  $\dot{x} = -y - z$ ;  $\dot{y} = x + 0.398y$ ;  $\dot{z} = 2 + z(x - 4)$

development as shown in Figure 7. We only need still to fold the two diverging 'wings' back on top of each other and bend the whole structure around a cylinder in order to obtain a picture of the structure of a simple chaotic attractor (see Figure 14. It shows the simplest known chaotic attractor - the so-called Rössler attractor).

A chaotic attractor can according to Thompson and Stewart (2002) briefly be described as 'a stable structure of long-term trajectories in a bounded region of phase space, which folds the bundle of trajectories back onto themselves, resulting in mixing and divergence of nearby states'.

When chaos develops in a nonlinear dynamic system - in a real life situation or a numerical simulation - the very first indication is a seemingly erratically oscillating behaviour of one or more components of the system. A hallmark of chaotic motion is its sensitivity to changes in the initial conditions, but there exist other indicators. According to Moon (1992) the following features characterize chaos:

1. Broad spectrum of the Fourier transform when the motion is generated by a single frequency (measured by fast Fourier transform (FFT))
2. Fractal properties of the motion in phase space, which denote a strange attractor (measured by Poincaré maps or fractal dimensions)
3. The complexity of the motion (measured by Lyapunov exponents). See a book on Nonlinear Dynamics e.g. Guckenheimer and Holmes (1983), Jordan and Smith (1988), Marek and Schreiber (1991), Moon (1992), Nayfeh and Balachandran (1995), Ott (1993), Thompson and Stewart (2002), Wiggins (1988) or Wiggins (1990).



**Figure 15.** A Poincaré section plot of one of Ueda's chaotic attractors.

When chaotic motion is suspected to occur, the first test is to watch a time series of a component of the motion. Then look at a phase diagram and at least one Poincaré section (see Figure 15). Next - if possible - repeat the process with slightly changed initial conditions. If the values of the state variables are changed in regions of the long time evolution when the initial conditions are changed, then the process is chaotic. You can also quantify the degree of disorder by computations of measures of the fractal dimension of the attractor or - the author's own favourite - a calculation of the largest Lyapunov exponent(s). If a series of calculations of the largest Lyapunov exponent converges to a positive value then the motion is chaotic.

When the dynamics is chaotic and multiple attractors exist, then the boundaries of the basins of attraction or the basins themselves may be *fractal dimensional*. As a consequence thereof there exist points  $\mathbf{P}_n$  in one of the basins of attraction with the property that there exists at least one point in the other basin of attraction in any neighbourhood of  $\mathbf{P}_n$ . In other words: When the boundary of a domain of attraction or the domain itself is fractal dimensional, then the result of an experimental investigation of the dynamic problem is in general unpredictable.

We shall conclude this brief introduction to chaos by defining *transient chaos*. It is a motion that looks chaotic and appears to show the characteristic features of a strange attractor, but it eventually settles into a regular motion or a chaotic attractor. It often appears when competing attractors exist in phase space. Chaotic transients must be treated with caution. They often decay with ever decreasing oscillations into a regular motion, but cases have been found where chaotic transients have amplitude jumps, which are larger than the amplitudes of the attractor, that the transient approaches. Such chaotic transients can be the cause of catastrophic developments in a dynamic system.

## 4 Parameter dependent Systems and Bifurcations

We shall limit the survey to systems with only one varying parameter  $\lambda$ . In vehicle dynamic problems the speed of the vehicle,  $V$ , is such a parameter. The solutions  $\mathbf{x}(t; \lambda)$  of the nonlinear dynamic problem

$$\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}; \lambda) \quad (4.1)$$

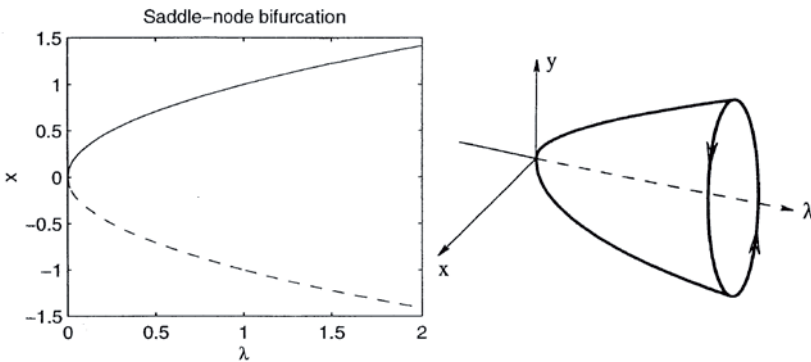
will then depend on the parameter  $\lambda$  as well as on time  $t$ . When we vary  $\lambda$  continuously, the solution will change continuously (according to our assumptions) and what is more important, *the number of solutions* may change i.e. in addition to the already known solution other solutions may appear or disappear at certain parameter values. Obviously our dynamic system is not structurally stable at such parameter values, since an infinitesimal change in the parameter value will change the phase diagram.

We investigate steady state solutions in this section. Their parameter dependence will be depicted in a *parameter-state space diagram*, where the parameter  $\lambda$  is the independent variable. In order to be able to present the diagram in a plane, the steady states in the state space are represented by a norm for instance the maximum amplitude or a Euclidean norm.

Very often in nonlinear dynamics a change in the number of steady state solutions happens through a *branching* or *bifurcation* from an existing solution under a continuous change of the parameter. Another common form of bifurcation is a solution that 'turns around' in the sense that the sign of the parameter increments changes in order to follow the solution in the parameter-state space diagram. Such a bifurcation is called a *fold* or *saddle-node* bifurcation, because the solution changes stability when 'it moves around the fold' (see figure 16). We call all these changes *local bifurcations*. They are very important for the analysis of the dynamics of the system, and they can only occur in nonlinear systems. They have in common that the parameter values where bifurcation occurs - *the bifurcation points* - can be found by an analysis of the parameter dependent eigenvalues of the associated Jacobian. We shall assume in the following that the steady state solution we want to investigate - the original solution - is *asymptotically stable* to start with. In all the cases the original solution will lose stability in the bifurcation point. This is the connection to the well-known linear stability theory.

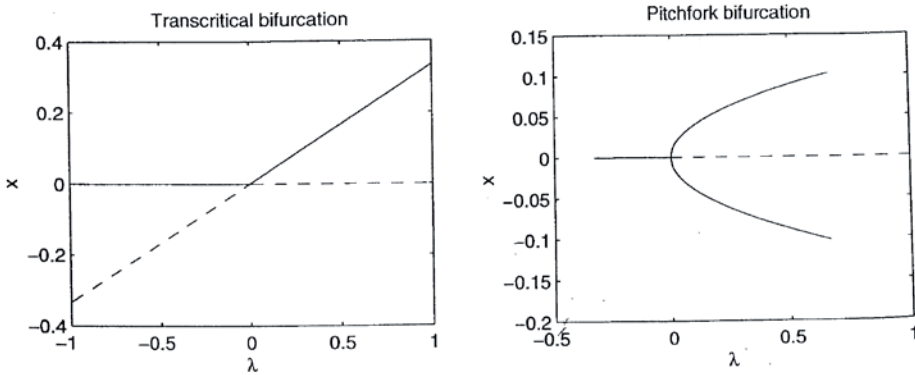
We can also introduce structural stability of bifurcations. The phrase, however, is hardly ever used in that context. Instead we denote a bifurcation *generic* i.e. 'commonly occurring' when a small change of the dynamic system does not change the properties of the bifurcation.

There are only two generic bifurcations from an equilibrium point under one control. They are the earlier mentioned *saddle-node*, when the largest eigenvalue of the Jacobian becomes positive, and the *Hopf bifurcation* (named after the mathematician, who first described it), which is a bifurcation of a *periodic solution* (see figure 16). In that case a complex conjugate pair of eigenvalues of the Jacobian crosses the imaginary axis in the complex plane into the positive real half plane. These two kinds of bifurcations are very common in vehicle dynamics.



**Figure 16.** A saddle-node bifurcation (left) and a Hopf bifurcation (right). A full line denotes a stable and a broken line an unstable solution. In the right diagram a stationary solution exists for all  $\lambda$ , but it is only stable for  $\lambda < 0$ . A stable periodic solution exists only for  $\lambda > 0$ .

If we allow more simultaneously acting controls then two additional generic bifurca-



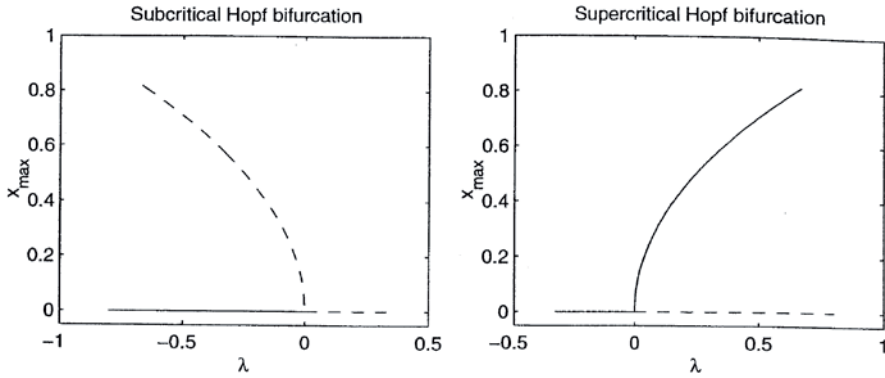
**Figure 17.** A transcritical bifurcation (left) and a pitchfork bifurcation (right). A full line denotes a stable and a broken line an unstable solution

tions are possible. They are the *transcritical bifurcation* and the *pitchfork bifurcation* or *symmetry breaking bifurcation*, when it occurs in systems that are symmetric in some way (see figure 17). In both cases the largest eigenvalue of the Jacobian becomes positive, and an additional analysis is necessary to distinguish them from the saddle-node bifurcation. Today *centre manifold reduction* and *normal forms* are commonly used for local bifurcation analyses. The centre manifold reduction theorem states that the change of dynamic behaviour of an  $N$ -dimensional dynamic system around the generic bifurcations can be described by the change of dynamic behaviour on a curve or two-dimensional surface in the parameter-state space. The curve or surface is called a *centre manifold* (see section 5). This important theorem guarantees that the bifurcations characterized by one or two eigenvalues of the Jacobian crossing the imaginary axis in the complex plane are the only generic ones that will be found in any  $N$ -dimensional dissipative dynamic system. The normal form is the simplest possible expression that is necessary to characterize the nonlinear dependence of the amplitudes of the solutions on the parameter near the bifurcation point.

There exist conditions for the stability of the bifurcating solution - or new branch - from equilibrium points. For the case of a Hopf bifurcation see figure 18. However, as a main rule an analysis of a bifurcation consists of three parts:

1. Determination of all the branches
2. Determination of the stability of the original solution on the other side of the bifurcation point
3. Determination of the stability of the new solution.

All these bifurcations are also generic for cycles under the same conditions regarding the control. A Hopf bifurcation from a cycle creates generically a quasi-periodic motion. Its trajectory lies on a torus in state space and the bifurcation is often called a *Naimark bifurcation*. For cycles there exists an additional bifurcation, the *flip bifurcation* or *period doubling bifurcation*. In a flip bifurcation a new periodic motion develops with twice the



**Figure 18.** A sub critical Hopf bifurcation, the bifurcating periodic solution is unstable (left) and a supercritical Hopf bifurcation, the bifurcating periodic solution is stable (right). A full line denotes an asymptotically stable and a broken line an unstable solution. In both cases the trivial solution loses its stability, when  $\lambda$  grows through 0

period of the original motion. A plot of the trajectory of a period-two solution could be the edge of a Möbius strip. The best known transition to chaos is an infinite sequence of period-doubling bifurcations.

For cycles the analogue of 'the eigenvalue of the Jacobian crossing into the positive real half plane' is 'the Floquet multiplier crossing the unit circle in the complex plane'. The Floquet multiplier in dynamic systems is the analogue of  $\rho_m$  introduced in section 3.3. If the Floquet multiplier crosses '1', then we have a fold, or a transcritical or a pitchfork bifurcation, where again additional calculations are necessary to determine the type. If the Floquet multiplier crosses '-1' we have a period-doubling bifurcation, and if it crosses almost anywhere else (a few special points are excluded) it is a Naimark bifurcation.

However there are other situations in which solutions can appear or disappear. Consider a case where we follow a certain solution in parameter-state space and a parameter value is reached at which a saddle-node appears somewhere in the parameter-state space away from the solution we follow. Relative to our 'certain solution' the saddle-node bifurcation is a *global bifurcation*. There exist many other kinds of global bifurcations, which due to their nature can not be investigated by 'local methods' such as a linearization around a known point in the parameter-state space. For further information the reader is referred to one of the several books on nonlinear dynamics e.g. Guckenheimer and Holmes (1983), Jordan and Smith (1988), Marek and Schreiber (1991), Moon (1992), Nayfeh and Balachandran (1995), Ott (1993), Thompson and Stewart (2002), Wiggins (1988) or Wiggins (1990).

For the sake of completeness we just mention *tori* as examples of other equilibrium sets in the state space. Quasi-periodic functions and multi-periodic functions all lie on tori in appropriate state spaces. We refer to the literature on nonlinear dynamics for information on their properties and the bifurcations that are connected with them.

## 5 The Centre Manifold Reduction Theorem

The Centre Manifold Reduction Theorem is the mathematical basis for linking the results we presented in the preceding section to vehicle dynamic systems with their in general high number of degrees of freedom. The theorem basically states that all the bifurcations that may be found in high dimensional parameter-state spaces in general are those described in section 4. The theorem is, however, proven only for systems of differential equations and is valid only for bifurcations from equilibrium points and periodic motion. Furthermore the dynamic systems must also satisfy certain conditions of smoothness, and vehicle dynamic systems will very rarely satisfy either one of these conditions. Vehicle dynamic systems are most often formulated as differential-algebraic systems, and the systems are rarely smooth. In the next section we shall therefore briefly describe some of the effects discontinuities like impacts and dry friction with stick-slip may have on the bifurcations.

## 6 On Non-Smooth Systems

New concepts that are important for the analysis of non-smooth systems in the state space will be introduced, and some of the possible bifurcations in non-smooth systems will be briefly surveyed in this section.

In non-smooth dynamic systems the location of the discontinuities in the state space must be defined through additional functions  $h_i(x_j) = 0$ ,  $i = 1, 2, \dots, M$ ;  $j = 1, 2, \dots, N$ , where  $M$  is the number of discontinuities and  $N$  is the dimension of the state space. In the  $N$ -dimensional state space the functions  $h_i(x_j) = 0$  define *hyper surfaces*, each of which splits the state space into two subspaces. The hyper surfaces are the so-called *switching boundaries* and they will in general intersect each other. Whenever a trajectory in the state space touches a switching boundary an *event* occurs. A simple example in the phase plane  $(x_1, x_2)$  is an elastic impact at  $x_1 = 0$ . When a trajectory in that phase plane hits the line  $x_1 = 0$  in the point  $(0, x_2)$ , it is the event, the velocity  $x_2$  changes sign and the trajectory will continue from the point  $(0, -x_2)$ . The outgoing trajectory will return into the half plane it came from with a slope, which has the opposite sign of that of the incoming trajectory. In the following section on numerical analysis we shall outline a numerical procedure for the handling of the events. Further material is found in Leine and Nijmeijer (2004, chap.6), where numerical integration methods for the solution of non-smooth dynamic systems are presented.

Systems that have a continuous as well as a discrete nature are also called *hybrid systems*.

In nonlinear, non-smooth dynamic systems the eigenvalues of the linearized system may cross the imaginary axis in continuous dependence on the control parameter, but they may also jump across the imaginary axis from the negative to the positive real half plane, when an event occurs. In case the eigenvalues depend continuously on the control parameter, the bifurcations are created in the way we already described in section 4, but the bifurcating solutions may depend differently on the control parameter in a neighbourhood of the bifurcation point. Slivsgaard (1995) found, in her analysis of the dynamics of the prototype single-axle bogie for the Copenhagen S-train that the

bifurcating stable, periodic solution will initially grow linearly with a finite slope from the bifurcation point. True (1994b) demonstrated in a simple example, that will be presented later, that the new behaviour is generic for a group of non-smooth dynamic systems.

In their book Leine and Nijmeijer (2004) discuss bifurcations of nonlinear, dynamic systems, where the eigenvalues jump across the imaginary axis. They call the phenomenon a *discontinuous bifurcation*. When the eigenvalues of the linearized system depend continuously on a single control parameter in the neighbourhood of the bifurcation point, then in the generic case only one real-valued or one pair of complex conjugate eigenvalues will cross the imaginary axis at a time. If however the eigenvalues jump in a discontinuous bifurcation, then more eigenvalues may jump at the same time, which may lead to different kinds of bifurcations. Some of the eigenvalues may even jump back again, so it is not at all easily recognized that they have jumped back and forth across the imaginary axis. Such cases may lead to new kinds of bifurcations. A similar situation occurs in maps. The analogue here is a jump of the Floquet multiplier(s) across the unit circle. We refer to Leine and Nijmeijer (2004) for more information.

Leine and Nijmeijer (2004) also discuss the bifurcations that result from smooth approximations of the non-smooth systems. They demonstrate in simple analytic examples that smoothing may change the character of the bifurcations. This feature revives the discussion of 'to which degree vehicle dynamics can be non-smooth'. In an impact contact forces are produced, but should the impact be modelled as an ideally elastic one with an instantaneous infinitely large contact force - possibly with a coefficient of restitution - or will the impact be modelled better by a dynamic law extending over a finite time interval? Both cases are non-analytic in the sense that discontinuities are introduced in the system, but the discontinuities are not equally strong. In the first case the velocity - the first derivative - has a finite jump; while in the second case the velocity varies continuously while the acceleration - the second derivative - jumps. The second case is an example of a weaker 'non-smoothness' than the first one. We shall not discuss modelling aspects here, but return to the topic in the section on numerical analysis.

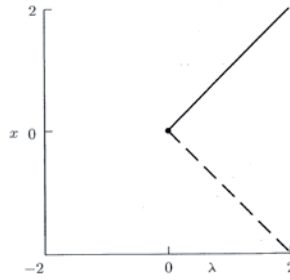
We continue this brief introduction with two simple examples of bifurcations of non-smooth systems. The first one is a non-smooth analogue of the saddle-node bifurcation in section 4. It is taken from Leine and Nijmeijer (2004). Let us consider the non-smooth system:

$$\dot{x} = \lambda - |x|. \quad (6.1)$$

There exist exactly two equilibrium solutions but only for  $\lambda > 0$ :  $x = -\lambda$  and  $x = \lambda$ . This yields the bifurcation diagram shown on Figure 19, which should be compared with Figure 16 (left).

When the two figures are compared, we find that the stable and unstable solutions meet at a finite angle in the bifurcation point, when the system is non-smooth, and the solutions depend linearly on the control parameter  $\lambda$ . This is in contrast with the classical saddle-node bifurcation on Figure 16 (left). A stability analysis of the two branches shows that the stability properties are the same in the smooth and the non-smooth cases.





**Figure 19.** A discontinuous saddle-node bifurcation. A full line denotes a stable and a broken line an unstable solution

The other example is a non-smooth analogue of the classical Hopf bifurcation shown on Figure 16 (right) in section 4. True (1994b) considered the system:

$$\dot{x} = y \tag{6.2}$$

$$\dot{y} = -(k\sqrt{x^2 + y^2} - \lambda)y - x, \quad k \neq 0. \tag{6.3}$$

We introduce polar coordinates by the transformation

$$x = r\cos v, \quad y = r\sin v$$

and obtain:

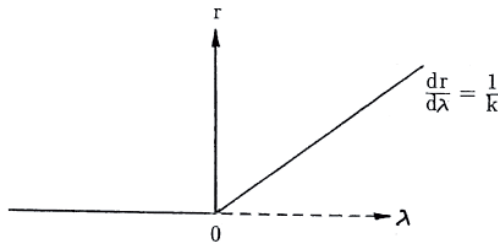
$$\dot{r} = (\lambda - kr)r\sin^2 v \tag{6.4}$$

$$\dot{v} = (\lambda - kr)\cos v \sin v - 1 \tag{6.5}$$

This system has the steady state solution:

$$\lambda = kr$$

in addition to  $r = 0$ .



**Figure 20.** A discontinuous Hopf bifurcation

A simple bifurcation analysis yields the bifurcation diagram on Figure 20. Here we see that the bifurcating periodic solution branches off linearly from the stationary solution under a finite angle in contrast to the classical Hopf bifurcation. A stability analysis of the three branches shows - like in the preceding example - that the stability properties are the same in the smooth and the non-smooth cases. In both cases the asymptotically stable stationary solution becomes unstable on the right hand side of the bifurcation point. The periodic solution is asymptotically stable, when it exists for positive values of the parameter  $\lambda$ , and otherwise it is unstable.

Impacts and stick-slip caused by dry friction contact are the two sources of non-smoothness that occur most often in vehicle dynamics. In nonlinear dynamic systems repeated impacts are considered to generate chaos in general. There is a lot of evidence hereof as well from experiments as from numerical simulations. When chaos is generated by repeated impacts it is important to bear the nature of chaos in mind. An infinitesimal disturbance from a trajectory on a chaotic attractor will develop into a large departure from the values on the original attractor at a later point in time. This is the well-known effect of the sensitivity on initial conditions in chaos. The departed value will, however, lie on another trajectory on the same chaotic attractor. Since the uncertainties of measurements as well as the error bounds in numerical tests are finite, it means that in reality it is impossible to follow a trajectory on a chaotic attractor. The best we can hope for in any case is to shadow the chaotic attractor by careful measurements or suitable numerical algorithms. The modelling of the impact itself will therefore have a negligible effect on the general dynamics after the impact - as long as the dynamics stays on the same chaotic attractor. We shall return to this subject in the following section on numerics.

This conclusion is also true for modelling of the stick in dry friction, but greater care must be exercised in that case. The stick has other important consequences for the dynamics. Stick will break some symmetries in the system. There is no proof of this, but the author has found this to be the case in all the vehicle dynamic systems his co-workers and he has analysed. Thereby chaos is more likely to develop. Stick may also change stationary points in the state space into *stationary domains* and more generally a steady state into a *continuum of steady states*. As an example of a stationary domain let us consider a rolling railway wheel set that carries a car body in a suspension with dry friction damping. On straight track the damper(s) may stick in positions that hold the wheel set fixed in a non-zero lateral displacement. The state is stable. The entire set of stable stationary lateral displacements and stationary yaw angles depends in general on the speed of the vehicle. In the case of the single rolling wheel set the stable stationary domains will constitute a three-dimensional hyper surface in the parameter-state space. The theory of bifurcations for smooth systems must be modified for dynamic systems with these properties. Such a wheel set may however, under different conditions oscillate and reach equilibrium on one of several different cycles in the state space, each of which depends on the initial condition. These cycles in the parameter-state space constitute a domain of steady states.

Leine and Nijmeijer (2004) also discuss dry friction problems. Other examples of the special effects of dry friction damping on the dynamics of vehicles may be found in the

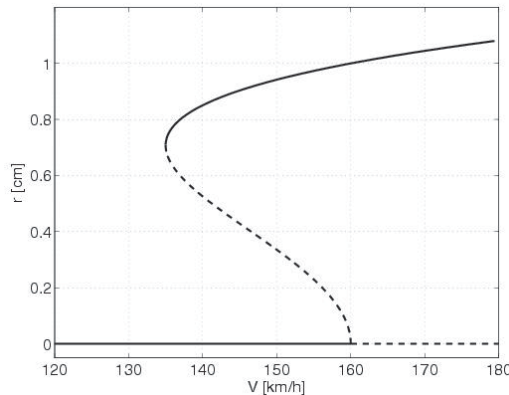
following chapter.

Leine and Nijmeijer (2004) write two chapters with 'Discussion and Conclusions'. We shall try to condense the conclusions relating to bifurcations that are most important for vehicle dynamics. For a deeper discussion the reader is referred to Leine and Nijmeijer (2004) or to the review paper by di Bernardo et al. (2007 ?).

There exist types of bifurcations in non-smooth dynamic systems that do not have a smooth counterpart. The mathematical theorems on bifurcations are almost all based on assumptions of some smoothness. Therefore *conclusions about the existence of bifurcations must be drawn from an inspection of bifurcation diagrams*. Unfortunately we are in general unable to find the unstable solution branches numerically; therefore the computed bifurcation diagrams will be incomplete. Leine and Nijmeijer (2004) analyse examples of bifurcations from a stationary solution in non-smooth systems. They find by analytical calculations that an eigenvalue of the Jacobian has crossed the imaginary axis. There exists, however, no mathematical proof that a crossing is a necessary condition for a bifurcation in a general non-smooth system. In our own numerical investigations of vehicle dynamic systems we found that the crossings of eigenvalues that are known from smooth systems also take place in our non-smooth systems. We found it to be true for bifurcations from stationary solutions as well as from limit cycles. See the review paper by di Bernardo et al. (2007 ?) for more information.

The conclusion is that *it is necessary to calculate bifurcation diagrams in order to determine the bifurcation points in vehicle dynamic problems*. This leads to the section 8 in which a simple numerical method for the determination of the *stable branches* of a general nonlinear dynamic problem will be described.

## 7 Dynamics of Railway Vehicles



**Figure 21.** A typical bifurcation diagram  $r$  versus the speed,  $V$ , for a railway vehicle. The full line indicates the asymptotically stable steady solutions and the broken line indicates the unstable steady solutions.  $r$  is a measure of the periodic motion e.g. the maximum lateral displacement of the front wheelset

On figure 21 is shown a bifurcation diagram for a theoretical problem. It has, however all the features that are typical for most conventional railway vehicles. It illustrates the dynamics of a passenger car designed for a maximum operational speed of 120 km/h running on a straight, ideal track.

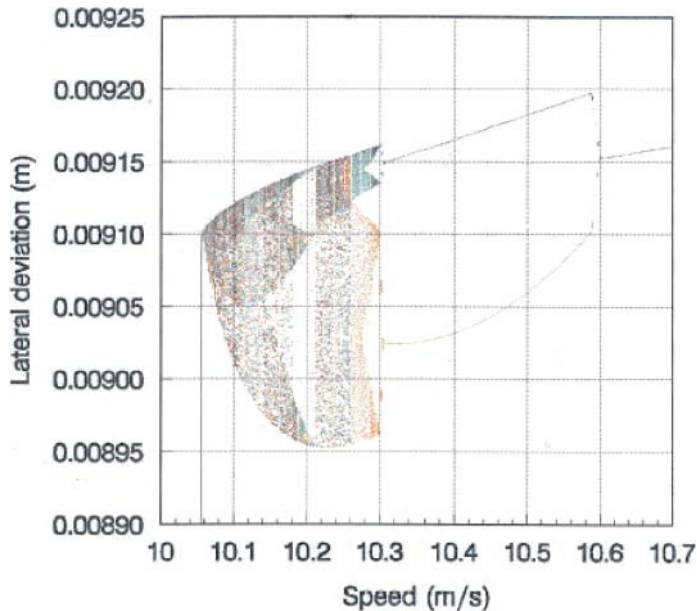
The diagram shows that the stationary solution for a growing speed will lose its stability at 160 km/h in a sub critical Hopf bifurcation. In this bifurcation a periodic solution with  $r > 0$  bifurcates away from the stationary solution. A stability analysis shows that the bifurcating periodic solution is unstable in the interval  $133.4 < V < 160$  km/h. At  $V = 133.4$  km/h the branch turns around towards higher speeds and gains asymptotic stability in a saddle-node bifurcation.

Notice that the stationary steady solution is unique below  $V = 133.4$  km/h. It is asymptotically stable for all speeds up to 160 km/h, but it is no longer unique above  $V = 133.4$  km/h. In the interval  $133.4 < V < 160$  km/h three steady solutions exist: one asymptotically stable stationary, one asymptotically stable periodic and one unstable periodic solution. Only the two stable solutions can be verified experimentally. Which one that will be seen depends on the track disturbances. If we start on the stationary solution in the interval  $133.4 < V < 160$  km/h and disturb it with a sufficiently large initial value, then the transient will end on the stable periodic solution. If the disturbance is sufficiently small the transient will approach the stationary solution. This is the typical situation in railway dynamics in this speed interval. The stationary motion is asymptotically stable to sufficiently small disturbances and unstable to sufficiently large disturbances. Since  $V = 133.4$  km/h in our example is the lowest speed at which a disturbance can change the motion to hunting, then it is natural to name that value of the speed *the critical speed*. Many tests have verified that this theoretical value corresponds very well to the value for the critical speed measured in real tests. The reason is that the disturbances on a normal track in daily use are finite and sufficiently large to 'throw' the stationary motion up on the periodic branch at - or at  $V$  a little higher than - the theoretical critical speed. The stationary motion loses stability at the higher speed  $V = 160$  km/h - the speed, which a stability analysis would yield. Above that speed the hunting mode is the only asymptotically stable solution in our example. The difference between the critical speed and the speed at which the stationary motion loses its stability is significant. It is very difficult to 'throw' a hunting motion back to the stationary solution. The reason is that the amplitude of the hunting usually is larger than the track disturbances. A disturbance must be of the right magnitude and fit accurately into the phase of the oscillation in order to silence the hunting, but a moment later another disturbance will 'throw' the mode back into hunting. The author has observed that phenomenon only once: During a test of an American gondola wagon on the TTCI test ring in Colorado. Hunting is thus more robust than the stationary mode in the speed interval with multiple attractors. In order to stop the hunting effectively the vehicle must slow down to speeds below  $V = 133.4$  km/h in our case. This hysteresis phenomenon is well documented in tests on real railway lines.

There are however other kinds of bifurcations possible. Slivsgaard and True (1994) investigated the dynamics of a simple model of a prototype single-axle bogie. It was used to test the steered single-axle bogies for the Copenhagen S-trains. The model wheel profile was conical with a conicity of  $1/20$  and the rail profile was an arc of a circle. The

flange was modelled by a very stiff spring with a clearance of 9.1 mm. The dynamic model is thus non-smooth, but it does not contain any algebraic constraint equations due to the linear wheel-rail contact geometry. A bifurcation diagram is shown on figure 22.

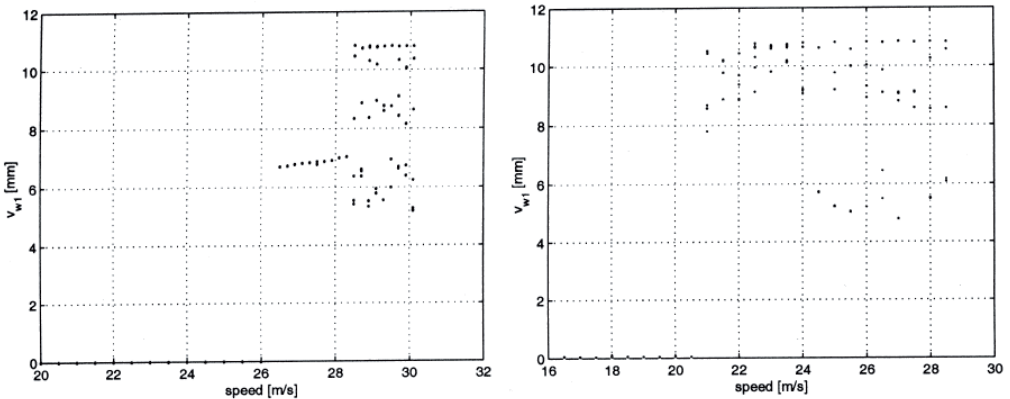
Now the bifurcation of the periodic motion is supercritical, and it is asymptotically stable. The creep/creep force law by Shen et al. (1984) is applied. Due to the second order discontinuity of the form  $x|x|$  in the SHE formula the bifurcation is of the type shown on figure 20. The motion turns chaotic in a *grazing bifurcation* (see di Bernardo et al. (2007 ?)) as soon as the lateral oscillations grow so large that the 'flange spring' comes into action. It is a type of bifurcation that only exists in non-smooth systems. In this special case a stability analysis of the stationary solution will furnish the critical speed, but we only know that for sure when a complete bifurcation analysis has been performed. The chaos only exists in a narrow speed interval before the oscillation becomes periodic and symmetric about the centre line of the track. The dynamics is best understood if the bifurcation diagram is described for decreasing speed starting above  $V = 10.6$  m/s, where the oscillation of the wheel set is periodic and symmetric. Just below  $V = 10.6$  m/s a symmetry breaking bifurcation splits the periodic motion into two asymmetric periodic motions that are reflections of each other with respect to the track centre line. The 'left hand' attractor is indicated by green colour on figure 22, and the 'right hand' attractor is red. The positive amplitudes are depicted. Close to  $V = 10.3$  m/s both periodic



**Figure 22.** Bifurcation diagram of a single-axle bogie. Its maximum lateral displacement versus the speed. From Slivsgaard and True (1994)

attractors start on what looks like an incomplete series of period doubling bifurcations that end in chaos when the oscillations 'graze' the stiff spring at the amplitude of 9.1 mm. An *inverse merger attractor crisis* (see Nayfeh and Balachandran (1995) s. 345) at  $V = 10.275$  m/s doubles the size of each of the chaotic attractors. At  $V = 10.260$  m/s the two asymmetric attractors merge in another inverse merger attractor crisis. It is clearly indicated on figure 22 by the mixture of the red and green dots. For still lower speeds one chaotic attractor exists, which is interrupted by the usual periodic windows. The attractor decreases in size with decreasing speed and vanishes when the oscillation of the wheel set stops hitting the 'flange spring'.

The next example shows that a supercritical bifurcation of the periodic solution from the stable stationary solution does not guarantee that other steady solutions do not exist at lower values of the speed than the Hopf bifurcation point. Xia and True (2003) investigated the complex dynamics of a Chinese gondola wagon on American three-piece-freight trucks. The bifurcation diagrams on figure 23 show the lateral displacement of the leading axle of the leading truck versus the speed.



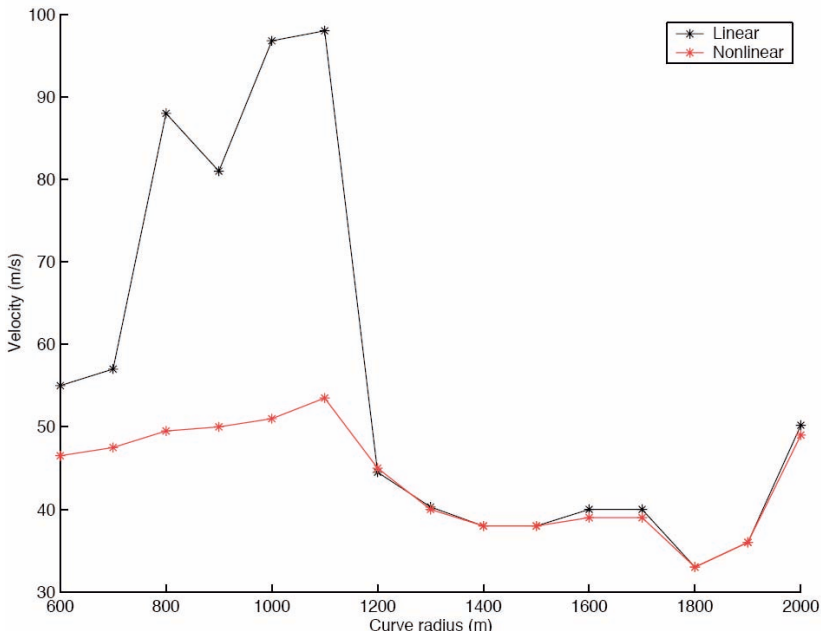
**Figure 23.** Bifurcation diagram of a Chinese gondola wagon. Only the asymptotically stable motions are shown versus the speed. On the left diagram the bifurcations for increasing speed are shown, and on the right diagram the bifurcations for decreasing speed are shown

On the left diagram a supercritical bifurcation into an asymptotically stable periodic motion is seen at  $V = 26.5$  m/s (95 km/h). This motion loses stability at  $V = 28.5$  m/s (102.5 km/h), and the only asymptotically stable motion found above 28.5 m/s is chaotic with large amplitude variations. On the right hand diagram it is seen that this chaotic motion prevails down to a speed of  $V = 21$  m/s (75.5 km/h) and remains asymptotically stable. The stationary solutions (due to the stick-slip there are infinitely many) are the only forms of motion only up to  $V = 21$  m/s. It is the critical speed in this case.

We may conclude that there exist two asymptotically stable modes in the interval  $21 < V < 26.5$  m/s. One is stationary and the other chaotic. In the interval  $26.5 < V < 28.5$  m/s there again exist two asymptotically stable modes - the same chaotic one and

a periodic motion. Above  $V = 28.5$  m/s only one asymptotically stable solution exists, and it is chaotic.

The plots on figure 23 were made in the following way: after each incremental change of the parameter,  $V$ , the solution of the initial value problem was computed in the time interval  $0 < t < 20$  s to eliminate the transient. Then the next maximum of the amplitude of the oscillation was recorded and plotted. The erratic distribution of the amplitudes is no proof of chaos but certainly an evidence of an aperiodic behaviour. Xia showed in his thesis that the dynamics is indeed chaotic.

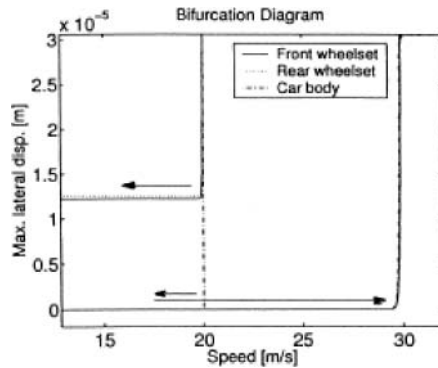


**Figure 24.** The critical speed (nonlinear) and the Hopf bifurcation (linear) versus the curve radius found by Petersen and Hoffmann (2002). The cant is  $2^\circ$ . Near  $R = 1200$  m a sub critical Hopf bifurcation suddenly changes to a supercritical Hopf bifurcation. Around  $R = 1800$  m the critical speed in the curve is lower than the critical speed on tangent track

All these theoretical results have in common that the track is assumed to be perfect and straight. Recent theoretical investigations by Zboinski and Dusza (2004), Petersen and Hoffmann (2002) and True et al. (2005) have demonstrated that the critical speed of bogie vehicles in curves of constant radius in general rather large radii will be lower than the critical speed on straight track. The radius of the curve must be so large that the vehicle is allowed to negotiate the curve at sufficiently high speeds. The critical speed depends also on the superelevation in the curve. On figure 24 a result from Petersen and Hoffmann (2002) is shown. Around a curve radius of 1800 m the critical speed of the bogie is lower than the critical speed of the bogie on straight track. The

lower critical speed in a curve has been observed in real life by Cooperrider and Law 1982 (personal communication).

Hoffmann et al. (2003) investigated numerically the dynamics of a two-axle freight wagon with UIC standard suspension. The UIC suspension carries the car body in links with rolling dry friction in the contact surfaces between the elements of the links. Hoffmann et al. (2003) have calculated a bifurcation diagram, which has a hysteresis loop. The calculation was started with symmetric initial conditions e.g. the lateral displacement of the wheel sets was zero. It is a stable stationary solution of the dynamic problem. After a supercritical Hopf bifurcation at  $V = 30$  m/s and going around in the hysteresis loop the stable stationary solution for the wheel sets did not return to zero at  $V = 20$  m/s. They both ended on a stable stationary non-zero solution. The result is shown on figure 25.

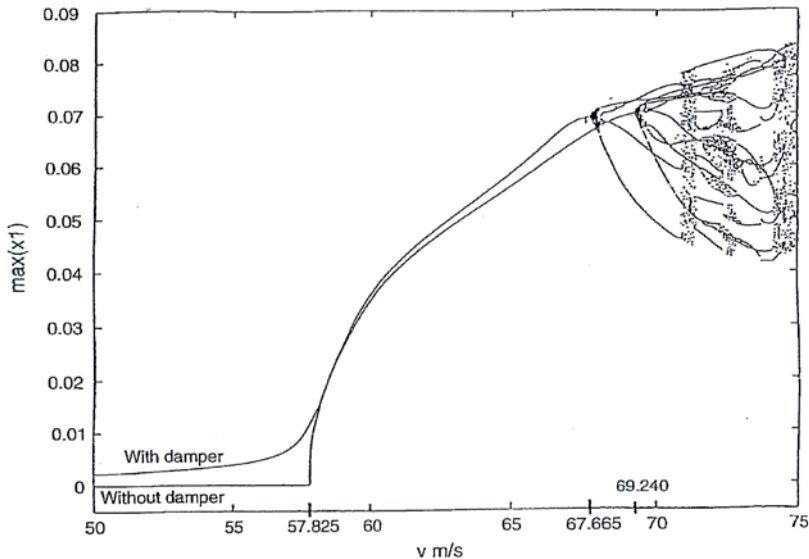


**Figure 25.** The 'open hysteresis loop' in a model of a freight wagon with dry friction in the suspension verifying the existence of a stable stationary domain

True and Asmund (2002) investigated numerically a model of a single wheelset under a freight wagon. The model was simplified to illustrate only the effect of the coupling between lateral dry friction dampers with stick-slip and the creep force. Therefore the kinematics of the wheel-rail geometry was linearized, and the wheels had no flanges. Figure 26 from True and Asmund (2002) shows an interesting comparison between the bifurcation diagrams for the wheelset without dampers and one with dry friction dampers. Without the dampers the dynamic system can be reduced to a set of differential equations, and the mathematical theory in chapter 4 applies. The bifurcation diagram is a classical one, with a trivial solution, which is asymptotically stable up to the Hopf bifurcation point, where a supercritical, asymptotically stable periodic solution branches off. The stationary trivial solution loses stability at the bifurcation point. For higher speeds the asymptotically stable periodic solution goes through a symmetry breaking bifurcation and a series of - what looks like - period doubling bifurcations, which lead to chaotic behaviour. With the dry friction damper the diagram is drastically changed. With the friction damper the trivial solution is no longer asymptotically stable. We now find a set of stable periodic motions with amplitudes, which grow slowly with the speed and



which depend on the initial conditions. There exists now a whole range of stable periodic solutions with different amplitudes. Only one is shown on figure 26. Its amplitude grows faster with the speed near the bifurcation point for the undamped bogie in a way that is known from nonlinear stochastic systems. The Hopf bifurcation has disappeared.



**Figure 26.** A comparison of the bifurcation diagrams for a single wheel set under a wagon without dampers and with dry friction dampers.  $\text{Max}(x_1)$  [m] versus  $V$  [m/s]

True and Trzepacz (2003) investigated the same wheelset, but they introduced a realistic wheel-rail geometry and dry friction yaw dampers in Asmund's model. The lateral component of the normal forces in the wheel-rail contact surface was ignored. For a large range of speeds the wheelset oscillated so much that the axle boxes hit their guidances. The oscillations were chaotic. The impacts were modelled as ideally elastic impacts, but in one series of simulations the impacts were modelled by linearly elastic dynamic systems, where the contact forces were calculated. The dynamics was hardly affected, but the computation time exploded.

Hoffmann (2006) finds in his work on the dynamics of two-axle freight wagons with the standard UIC suspension that the Hopf bifurcation from the stable set of stationary solutions is a classical one. This is in spite of the fact that the problem contains much non-smoothness. Hoffmann (2006) gives the following reason: The motion for small lateral and yaw displacements around each of the stationary states is smooth. For such small motions the links in the suspension roll on each other, and the suspension acts as a linear spring without stick and dissipation.

These examples emphasize that the dynamics of non-smooth systems must be investigated with great care!

## 8 On the Numerical Analysis of Vehicle Dynamic Systems

### 8.1 The Numerical Solution of Smooth Problems

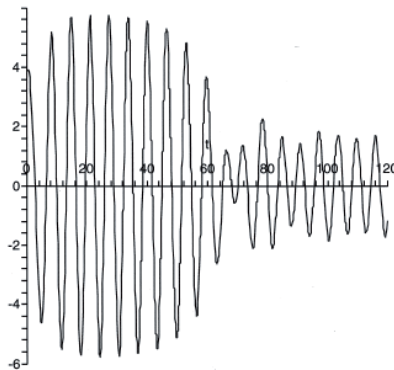
In this section we shall briefly survey the numerical solution of vehicle dynamic problems. Special attention will be paid to problems with coexisting attractors and to non-smooth problems. We shall demonstrate how such problems must be treated in order to obtain reliable dynamic results.

When coexisting attractors exist in a nonlinear dynamic problem, the steady state solutions of a specific initial value problem depend on the initial values. This is in contrast to linear dynamic problems, where coexisting attractors cannot exist. On the other hand the solution of a given nonlinear initial value problem is unique provided certain continuity conditions are satisfied. In other words: The trajectory in the state space is uniquely defined. If, however, two trajectories that converge towards two different attractors are sufficiently close in some part of the state space, then numerical time step integration along one of the attractors may jump to and follow the other attractor. This leads to an erroneous numerical solution of the initial value problem. The jump may be the result of a bad choice of error bounds or time steps, but it may also be the result of the use of the wrong numerical routine. We shall demonstrate below the erroneous result of the application of an Euler explicit routine (which should *never* be applied to nonlinear equations (!)) to a simple nonlinear initial value problem.

In section 2.2 we solved:

$$2.56(d^2x/dt^2) + 0.32(dx/dt) + x + 0.05x^3 = 2.5\cos t; \tag{8.1}$$

numerically with two pairs of initial conditions. Two different attractors were found, and they are shown on figure 6.



**Figure 27.** The numerical solution to the initial value problem (8.1)+(8.2) using a Runge-Kutta 45 routine. The transient approaches the correct attractor

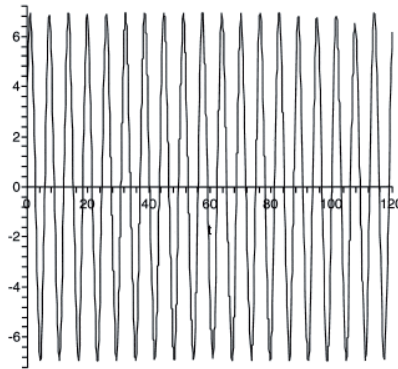
Let us now consider equation (8.1) with the initial conditions:

$$x(0) = 3.15, \quad dx/dt(0) = 5.81; \quad (8.2)$$

We first solve the initial value problem (8.1)+(8.2) with a Runge-Kutta 45 routine from MAPLE 9.5. The transient approaches the correct attractor with small amplitude that belongs to the initial conditions (8.2). The result is shown on figure 27.

We next solve the same initial value problem (8.1)+(8.2) with an explicit Euler routine - also from MAPLE 9.5. The result is shown on figure 28, and it is clearly seen that *the transient in this case approaches the other - and incorrect - attractor*.

The explicit Euler routine is very popular, because it works fast. As long as no other solution or measured results from a test exist for comparison, the Euler solution will be accepted. It is outright dangerous! The explicit Euler routine should never be used for the numerical solution of nonlinear dynamic problems.



**Figure 28.** The numerical solution to the initial value problem (8.1)+(8.2) using an explicit Euler routine. The transient approaches *the wrong attractor*

This simple example demonstrates that great care must be taken, when a numerical solver is chosen for a given nonlinear dynamic problem.

Garg and Dukkipati (1984) discuss and demonstrate the use of explicit and implicit routines for the numerical solution of linear as well as nonlinear dynamic problems. They test a selection of routines on both linear and nonlinear dynamic problems. They compare CPU time, stability and artificial damping in their examples. In an *explicit routine* the next step  $x_{n+1}$  depends explicitly on known quantities only. In the nonlinear example with the rolling wheel set Garg and Dukkipati (1984, p. 44) find that the explicit solvers are either unstable or introduce artificial damping. In *implicit routines* the dependence is implicit, so the next step has to be calculated by the solution of an equation. Since vehicle dynamic problems are nonlinear the said equation will also be nonlinear. It means that a nonlinear equation has to be solved numerically for each time step for example by a *predictor corrector method*. The explicit solvers are inefficient for the analysis of the so-called *stiff* problems. Generally a problem is said to be stiff, if the solution contains

both slow and very fast processes. The fast processes decay very fast with time  $t$ , so a very short step length is needed to keep the explicit solvers from 'blowing up' in the numerical solution. Furthermore the explicit solvers often become unstable when they are applied to stiff problems. In order to maintain stability the step size again needs to be very small. Vehicle dynamic problems as well as the numerical simulations of real world tests are usually stiff problems. Therefore the application of implicit routines is recommended for the numerical solution of these initial value problems.

Rill (2006) has applied an implicit Euler routine to vehicle dynamic problems with success. True and his co-workers have most often used the LSODA routine for the calculation of bifurcation diagrams - also successfully. LSODA changes solution strategy automatically depending on whether the problem is stiff or not. The Runge-Kutta solver rk45 has been recommended for years as the most effective routine for the purpose of shadowing chaotic attractors, but it is an explicit routine, and it therefore cannot be recommended for the numerical analysis of vehicle dynamic problems. There exist, however, also implicit Runge-Kutta solvers for stiff problems. All these Runge-Kutta solvers are a part of MATLAB. Most recently the routine SDIRK by Nørsett and Thomsen (1984), based on the theory from Alexander (2003), has been applied by Hoffmann (2006) to the numerical analysis of two-axle freight wagons - a stiff problem with rolling and sliding dry friction and bumper stops. SDIRK is a recent implicit routine, which can be downloaded from the web. For a detailed description of time-stepping explicit and implicit numerical routines and recommendations for their applications the interested reader is referred to the literature Lambert (1972) and Öderlind (1998).

The time-stepping numerical methods are by far the most popular numerical methods for the solution of vehicle dynamic problems, because they lend themselves naturally to the solution of initial value problems. It should be mentioned, however, that steady state solutions have been found by numerical solution of the related optimization problem. It is claimed (oral information) that the stable attractors in a bifurcation diagram for a vehicle can be computed faster in that way.

The engineer without a good background in numerical analysis faces a problem when he must choose a differential equation solver (DE-solver). Although a numerical solver may deliver a solution when it is applied to a given dynamic problem, it cannot be taken for granted that the solution is the correct one! It is strongly recommended that the engineer applies two different and acknowledged routines to the same problem - at least initially. If the result of the calculations using one of the solvers differs considerably from the results, which the other solver yields, he must conclude that at least one of the solvers is bad in his case. On the other hand he cannot be sure that the results are correct even if the two results agree with each other!

## 8.2 The Numerical Solution across a Discontinuity

When a vehicle dynamic problem is not smooth, the dynamic problem must be supplemented with functions, which define the location of the discontinuities in the state space:  $h_i(x_j) = 0$ ,  $i = 1, 2, \dots, M$ ;  $j = 1, 2, \dots, N$ , where  $M$  is the number of discontinuities and  $N$  is the dimension of the state space. The functions are *hyper surfaces* in the state space, see chapter 6. By the word 'discontinuities' is understood not only discontinuities

in the functions themselves (zero-order discontinuities) but also discontinuities in the derivatives up to some order (higher-order discontinuities). The hyper surfaces are the so-called *switching boundaries*. As an example we consider now a dynamic problem with one switching boundary  $h(x_j) = 0$ . The analysis is easily generalized to problems with more switching boundaries as long as we can deal with them one at a time. In the two domains, divided by the switching boundary, the problem is smooth, and there we solve the initial value problems in the usual manner. We assume that  $h(x_j) < 0$  on one side of  $h(x_j) = 0$  and  $h(x_j) > 0$  on the other side. All we in addition need to consider is the region close to the point P, where the solution crosses from one domain to the other. The point P is called the *transition point*.

When we integrate our dynamic problem as an initial value problem we must stop the integration in the first domain at the switching boundary. The switching boundary is only defined in terms of the *dependent variables*  $x_j(t)$ . We must therefore continuously check in our solution procedure at which time  $t_0$  the trajectory is within a given very small distance from  $h(x_j) = 0$  or has already crossed the switching boundary. It is then necessary to reformulate the solution procedure to determine the time T iteratively at which the trajectory hits or touches the switching boundary. T can only be found numerically within a certain tolerance, but it should be determined as accurately as possible. The trajectory (dynamic problem) is then integrated up to the time T and stopped. We have now found the transition point P. It is defined by the values  $x_j(T)$ ,  $j = 1, 2, \dots, N$ .

In the mathematical formulation of the non-smooth dynamic problem it must be defined what will happen at the switching boundary. The collision with the switching boundary and what happens there is called an *event*. As an example of an event we mention a jump  $\Delta v$  in the first derivative of one of the state variables - say  $x_K$  - at time T.

We then formulate a new initial value problem at the time T using the end state variables found at the transition point P, adjusted by the law of the event, as the new initial values for the integration. In our example we take all our end state variables at the time T and only add the given jump  $\Delta v$  to the derivative of  $x_K$  at the time T. The continuation of our dynamic problem is then given by the new initial values at the time T combined with the smooth dynamic formulation of our problem in the *relevant domain*. In our example the relevant domain will be on the other side of the switching boundary - the second domain. Then the integration continues in that domain.

As an example we investigate the dynamics of a wheel set used on a railway wagon and presented in Garg and Dukkipati (1984, p. 44). The system is shown on figure 29, where the wheel set is travelling along an ideal track.

The nonlinear equations of motion for the lateral displacement and the yaw angle and the values of the different physical parameters can be found in Garg and Dukkipati (1984, section 2.4.2).

The variation of the difference in the rolling radii of the wheel set versus the lateral displacement is shown on figure 30.

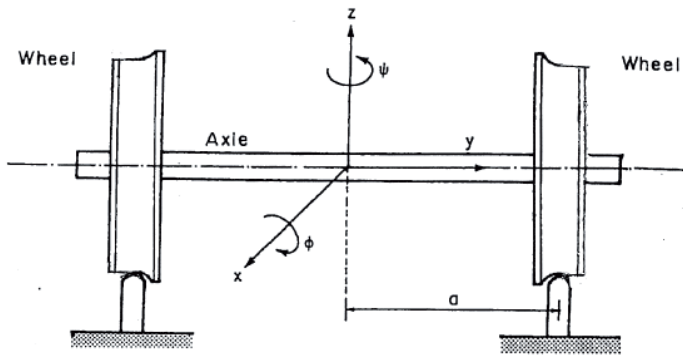


Figure 29. The wheel set model

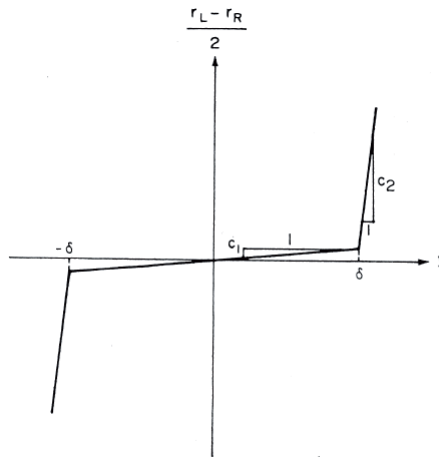
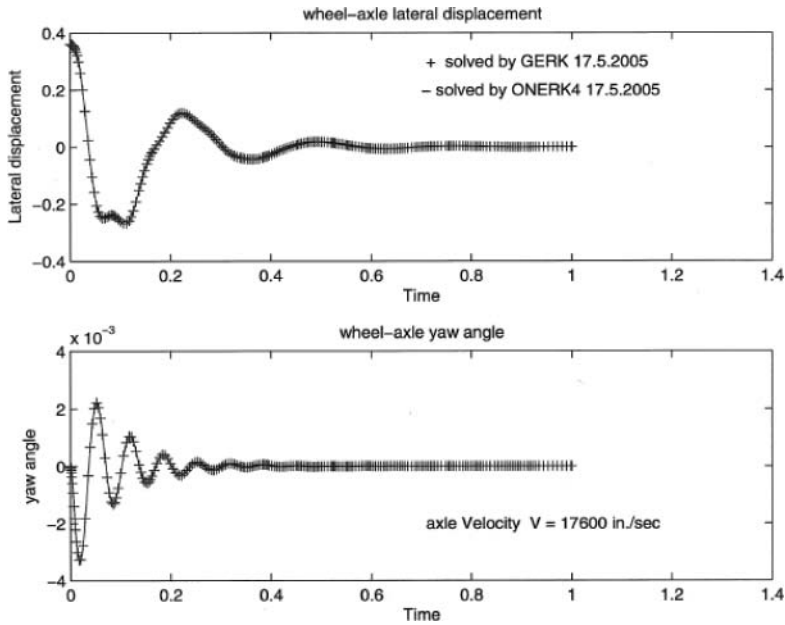


Figure 30. The variation of different variables with lateral displacements

For different entries we have the following values of the constants  $c_1$  and  $c_2$ ; the distance  $\delta = 0.35$  in the experiments.

var	$c_1$	$c_2$
$\Delta_1$	0	18
$\Delta_2$	0	18
$\Delta_L$	0.15	22
$\frac{r_L - r_R}{2}$	0.005	10



**Figure 31.** Wheel set dynamics solution for  $V = 160$  km/h

When we solve the problem for an initial lateral displacement  $y = 0.36$  with a speed of 160 km/h we obtain the solution on figure 31, where the damping is strong and only one passage of the discontinuity occurs. Here the solution with constant step size called ONERK4 and the solution that tracks down the discontinuity called GERK are close.

When the velocity is increased to near the Hopf bifurcation point of *this purely theoretical problem*, where more passages of discontinuities occur, we find very different solutions. The constant step size solution is much less accurate than the one where the discontinuities are tracked as seen on figure 32. This leads to different interpretations of the dynamic properties. It may be critical since this case is near the point where the stationary solution becomes unstable.

Let us summarize our experience gained from different numerical treatments of discontinuities in vehicle dynamic problems. The simplest way to handle the discontinuities is to *ignore the discontinuities and integrate across them as if they do not exist*. One hopes that the 'smoothing property' of the numerical routine will automatically take care of the problem with the discontinuity. If the discontinuity is in the second or higher derivatives of the functions in the dynamic problem this strategy may work well. The general attitude is that as long as the numerical integration delivers a solution then the solution is 'the correct one'. This is rarely true, as we just demonstrated in the last example. A solution found in this way may even be qualitatively wrong. A chaotic dynamic behaviour may be totally missed. This was illustrated by Xia and True (2003). Last but not least the numerical solver will consume a lot of CPU-time with the integration across the discontinuity, and we do not know the resulting numerical approximation of

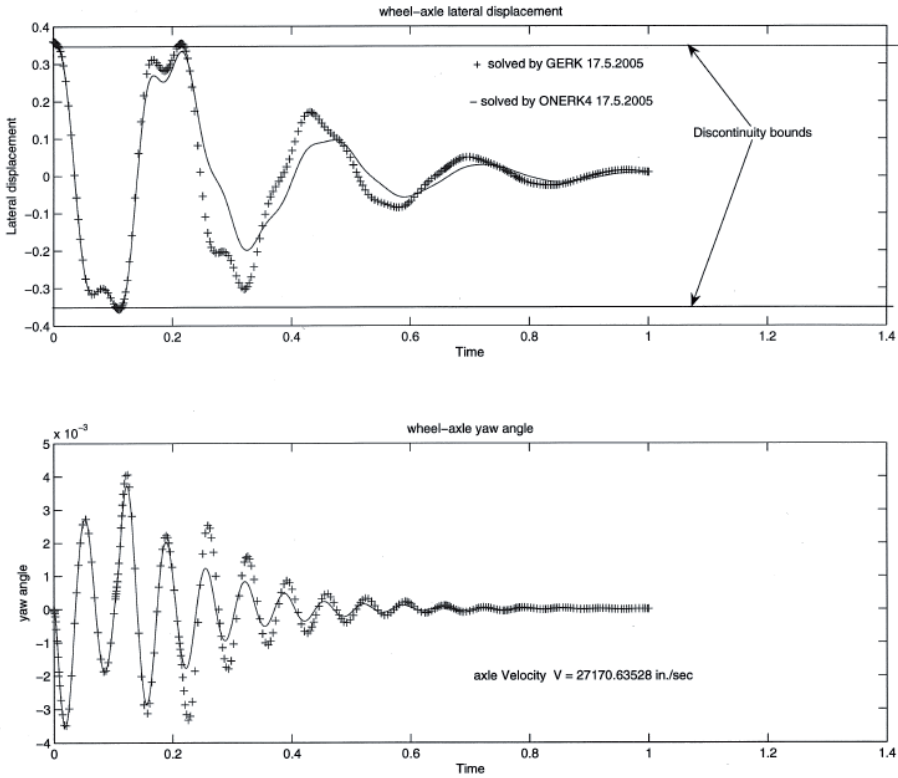


Figure 32. Wheel set dynamics solution for  $V = 2484 \text{ km/h}$ , a purely theoretical problem

the smoothing function.

The problem with the resulting smoothing function leads to the second method of handling the discontinuities: *The smoothing functions are prescribed in each and every case.* Then - at least - we know how the numerical routine handles the discontinuity, and the CPU-time will be reduced. The smoothing function must model the physical reality in the best possible way. Many people claim that this method is the optimal one, because it models the physical reality closely. The problem with the method is, however, that 'a good mathematical model' is not always known. In the articles by True and Asmund (2002) and True and Trzepacz (2003) their own smooth model of stick-slip dry friction in plane contact between steel or cast iron surfaces was applied. The results of the numerical dynamic analysis contain a lot of chaotic dynamics as could be expected from the nature of the problem. A comparison was, however, made between an ideally kinematical model and a dynamic model of an *elastic impact*. When the moving part had 'entered' the part in rest the velocity of the moving part changed sign in the kinematical model. In the dynamic model the law of elastic deformation was applied. True and Trzepacz (2003) used a constant and very small steplength in the integrations



of the dynamic model. The application stopped, when the two parts were separated again. Only the dynamic model is able to yield an estimate of the contact forces. In the dynamic model the CPU-time used for integration across the impact exploded. The results differed - as could be expected - but both results lie on the same chaotic attractor. It is interesting to note, however, that the dynamics of the system is the same chaotic one in the two cases. The situation was discussed in chapter 6 with relation to the modelling of the impact. The robust dynamics is closely connected with the chaotic dynamics in this case. The robustness is, however, not guaranteed. In other problems the dynamics of the two cases may differ.

The third - and correct - way of handling the discontinuities numerically is to *track down the transition point as accurately as possible*. The transition point is the point in the state space where the trajectory hits the switching boundary. *Then stop the time integration of the dynamic problem at that point, apply the law of transition and continue the time integration of the dynamic problem on the relevant side of the switching boundary*. This is the safest and most accurate method to deal with non-smooth problems. The extra CPU-time spent in the determination of the transition point is much less than the time not spent for the integration of the dynamic problem with very small time steps along steep gradients across the discontinuity. It is the logical numerical application of the results presented in section 6. The method was used with success by Hoffmann (2006) to investigate the difficult problem of the dynamics of two-axle freight wagons. His models contain both rolling and sliding friction in the suspension as well as possible impacts between the axle boxes and their guidance. The lack of proper consideration of the discontinuities in the numerical analysis leads in general to an even qualitatively erroneous result of the calculations.

It is important to identify coexisting attractors in nonlinear dynamic systems. The necessity is obvious, when the attractors are characterized by different amplitudes. It may seem less obvious if the attractors have almost the same amplitudes and are only qualitatively different. If however - like in True and Asmund (2002) True and Trzepacz (2003) and Xia and True (2003) - one of the coexisting attractors is a chaotic attractor, then great caution must be exercised! We refer to the discussion of chaotic transients at the end of chapter 3. The existence of a coexisting chaotic attractor will often give rise to violently oscillating transient motions with dangerously high amplitudes. This happens for transients approaching the chaotic attractor as well as for transients approaching a non-chaotic attractor. It is therefore very important to identify all the chaotic attractors in a nonlinear dynamic problem. The possible large amplitude transients may lead to accidents in the physical world.

### 8.3 A way to compute stable branches of vehicle dynamic problems numerically

In order to find the stable branches - the attractors - of a vehicle dynamic problem, use *path following in state parameter space*. We shall assume that *the track is ideal and free of disturbances*. Remember that we want to find the 'eigendynamics' of the vehicle. We have given a stable steady state solution  $\mathbf{X}(t)$  to our dynamic problem for a certain value  $V_0$  of the control parameter, which in our case is the speed,  $V$  (see figure 21). We

can then find the vector solution at a time  $t_0$  and store the state vector  $\mathbf{X}(t_0)$ . It is called the end value. The path following is performed stepwise. In each step we increase or decrease  $V$  with a small increment and calculate the steady state solution of the resulting dynamic problem using the end value of the preceding step as the initial value for the new step in  $V$ . We take as many steps as needed. The process will end, when the transient in the integration approaches another stable steady state solution  $\mathbf{X}_1(t)$  for a certain value  $V_1$  of the parameter  $V$ .  $V_1$  is close to a bifurcation point that has just been passed, and which can now be determined by interpolation of the parameter  $V$  or by the Newton-Raphson method in order to find the bifurcation point  $V_c$ . In this way a solution of the dynamic problem can be traced or *followed* and bifurcation points found in the parameter-state space. At the bifurcation points we turn our attention to the new stable branches and follow them in dependence on  $V$  for both growing and falling values. Apply as the initial condition the end values of the new steady state solution  $\mathbf{X}_1(t)$ . It is highly recommendable to calculate the eigenvalues of the dynamic problem, linearized around the stable steady state solution and find their change with the parameter across the bifurcation point assuming that the Jacobian exists. The calculation is a valuable verification of the properties of the found bifurcation. In a bifurcation of a periodic solution from a stable stationary solution two complex conjugate eigenvalues should cross the imaginary axis at the bifurcation point from the negative to the positive real half plane.

The success of the method is based on the assumption that the stable branches can be found as bifurcations from at least one stable steady state solution (branch) of the problem that is known a priori. It must be emphasized that it is not guaranteed that all the stable branches can be found in this way in all dynamic problems. In vehicle dynamics the aim is to design a vehicle that runs smoothly along a given track. We shall therefore make the basic assumption that the vehicle dynamic model to be investigated has a known steady state solution in a certain speed interval in an appropriate moving coordinate system. In vehicle dynamics the speed,  $V$ , is an appropriate control parameter, and in most cases 'the appropriate coordinate system' is one that moves with the constant speed,  $V > 0$  along the track of the vehicle. A dynamic model of a 'good' vehicle will most often have a trivial stable solution for sufficiently small values of the speed  $V > 0$ , i.e.  $\mathbf{x}(t) = 0$ . It can be used as the initial solution.

The path following procedure can be accelerated and automated by application of the *ramping method*. In the dynamic system the control parameter  $V$  is replaced by a monotonically slowly growing or decaying function of time,  $V(t)$ . In the ramping method we start the integration of the dynamic problem as usual with the known end value and integrate up to an appropriate end time  $T$ , so that  $[V(0), V(T)]$  covers the designed speed range of the vehicle. Repeat the procedure as long as necessary to determine all the attractors. The bifurcation diagram will in practice be quite well approximated by the ramping solution, but the method overshoots the bifurcation points, and one or more of closely situated bifurcation points may be missed. The bifurcation points must afterwards be recalculated with greater accuracy.

The smallest parameter value  $V_{crit}$ , for which there exists another stable steady state solution (attractor) to our dynamic problem in addition to the stationary one, is the *critical speed* of the vehicle. The path-following method has been used successfully not

only by True and his co-workers. It is described and discussed by True in several papers True (1994a, 1999, 2004), and its results have been verified in road tests.

There exist good commercial vehicle dynamic simulation routines. When a potential user must select a program, he must test his requirements of the program against the performance of the program and against the requirements listed in this section. The programs must contain at least two recognized implicit solvers for the solution of nonlinear dynamic problems, and the user must be able to control the numerical parameters such as step lengths - constant or varying size - and error bounds. Explicit Euler routines are of course banned from use!

If not treated properly the interpretation of the dynamic properties from the numerical results may be very different from the properties of the real system.

We recommend when the numerical tools are selected, that a similar care is spent as the care taken of the modelling. It is of the outmost importance that properties that influence the dynamic behaviour such as the discontinuities are properly accounted for.

## 9 A Guide to Vehicle System Dynamics

Due to the nonlinear coupling with the supporting ground it is basically impossible to split the dynamics of the total vehicle-ground system into effects caused by the properties of the vehicle and those of the support. The contact forces between the rails and wheels are important unknowns in the simulations not only because they act as an input to the dynamic vehicle model. They are important for the design of the vehicle and the deterioration of the structure of the support. We are however lucky in the sense that the dynamic influence of the vehicle on the ground acts on a much larger timescale than the instantaneous action the vehicle is exposed to from the ground. The feedback can therefore be neglected in the short-term modelling of the vehicle dynamics.

It is however possible to make deductions and achieve a better understanding of the dynamics of the total system if the dynamics of some of the elements in the system is known. The dynamic features of a vehicle model should therefore be investigated first. The investigation of the full nonlinear vehicle dynamic model starts with a numerical calculation of the bifurcation diagrams for the vehicle running on an ideal straight track (section 8.3). The calculations must be performed for all relevant contact geometries and for a relevant selection of adhesion coefficients. The speed is chosen as the control parameter, while all other parameters are kept constant. The relevant attractors in the parameter-state space must be found and the critical speed determined. If necessary the calculation of the bifurcation diagrams can be repeated with other parameter combinations.

Next the investigation continues with a set of numerical calculations of the bifurcation diagrams for quasi-stationary curving in arcs with selected radii. Again the speed is chosen as the control parameter. A lower critical speed will often be found in curves with a large radius than the one found by driving on tangent track.

When the attractors are known we are ready to make numerical simulations of the vehicle dynamics on realistic tracks, transition curves, turnouts and in other situations where the transient response is important. These simulations serve also as a basis for comparisons with measured results from road tests and as a verification of the dynamic

model. This may give rise to a great variety of nonlinear motions. In a master thesis Engbo Christiansen (2001) investigated the dynamics of the Cooperrider (1972) bogie on a sinuous track. Depending on the phase difference between and the amplitude of the sine functions of the right and left rail Engbo Christiansen (2001) found many of the nonlinear phenomena in the books - but only when the speed was above the critical speed. The dynamics above the critical speed was obviously a result of the mode interactions between the hunting motion of the vehicle and the excitations from the track. Since most railway freight vehicles today often run at supercritical speeds special care must be taken when their dynamics is modelled and investigated.

In every step of the investigation it is important to pay attention to the numerics. After having spent so much effort on the construction of the theoretical model many investigators seem to believe that the job is done. More effort is needed at that point than that spent by taking just any numerical solver, that can be found on the web or that is incorporated in a commercial simulation program - often without any information about the type of solver and its parameters. We recommend when the numerical tools are selected, that a similar care is spent as the care taken of the modelling. It is very important that properties such as the discontinuities which influence the dynamic behaviour are properly accounted for, see section 8.2.

It is of course permitted to make approximations of solutions to nonlinear systems for instance by linearization of nonlinear characteristics or terms in the dynamic system. If the total dynamic system shall be linearized, it is important that the operator knows which stable solution i.e. stable branch he will linearize around. He must also investigate whether a linearization is permitted - see section 3.2. The Jacobian  $\mathbf{J} = [\partial f_i / \partial x_j]$ ,  $i, j = 1, 2, \dots, N$ ; must exist, and the rest term must satisfy  $\|\mathbf{G}(\mathbf{x}; \boldsymbol{\lambda})\| / \|\mathbf{x}\| \rightarrow 0$  for  $\|\mathbf{x}\| \rightarrow 0$  if the trivial solution is considered. It is therefore important to know the full nonlinear operator as well as its bifurcation diagram first. The linearized operator can yield valuable information about certain properties of the vehicle model. The eigenvalues will for instance inform about characteristic frequencies and their attenuation and potential resonances in the vehicle dynamic system.

When the analyses of the nonlinear dynamics of vehicle systems are performed

- with accurate theoretical multibody system models that consist of elements with well defined and verified dynamics
- with due regard to the nonlinear nature of the systems
- on modern computer systems
- by well educated personnel using software that is carefully selected with regard to the nature of the specific dynamic problem,

then the results will be so accurate that simulations in future can be used as a documentation for certain vehicle performance requirements and even as a part of the safety specifications. This will give the manufacturers, the national boards of safety and large user organisations like the railway companies a strong and reliable tool for the design and testing of vehicle systems under realistic or even unrealistic operating conditions. The models may also be valuable tools for investigations of accidents. The advantages of such theoretical analyses are lower costs compared with alternative methods, flexibility

with respect to design parameters and operating conditions and - in many cases - a faster production of results.

## Bibliography

- R. Alexander. Design and implementation of DIRK integrators for stiff systems. *Applied Numerical Mathematics*, 46:1–17, 2003.
- J. Birkedal Nielsen. *New Developments in the Theory of Wheel/Rail Contact Mechanics*. PhD thesis, IMM, The Technical University of Denmark, 1998.
- A. Chudzikiewicz. *Selected Elements of the Contact Problems necessary for investigating the Rail Vehicle System Dynamics*, pages 273–304. Advanced Railway Vehicle System Dynamics. Wydawnictwa Naukowo-Techniczne, Warsaw, Poland, 1991.
- N.K. Cooperrider. The hunting behavior of conventional railway trucks. *ASME J. Engineering and Industry*, 94:752–762, 1972.
- S. Damme, U. Nackenhorst, A. Wetzel, and B. W. Zastrau. *On the numerical Analysis of the Wheel-Rail System in rolling Contact*, pages 155–174. System Dynamics and Long-Term Behaviour of Railway Vehicles, Track and Subgrade. Springer, Berlin, Heidelberg, 2003.
- M. di Bernardo, C. Budd, A.R. Champneys, P. Kowalczyk, A. Nordmark, G. Olivar, and P.T. Piiroinen. Bifurcations in non-smooth dynamical systems. *Submitted to SIAM Review*, page ?, 2007 ?
- L. Engbo Christiansen. The dynamics of a railway vehicle on a disturbed track. Master's thesis, Department of Physics, The Technical University of Denmark, 2001.
- V. K. Garg and R. V. Dukkipati. *Dynamics of Railway Vehicle Systems*. Academic Press, Toronto-Orlando-San Diego-New York-London-Montreal-Sydney-Tokyo, 1984.
- J. Guckenheimer and P. Holmes. *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*. Springer, Berlin, Heidelberg, 1983.
- H. Heumann. *Grundzüge der Führung der Schienenfahrzeuge*. Sonderdruck aus Elektrische Bahnen. Verlag R. Oldenburg, München, 1950-1953.
- M. Hoffmann. *Dynamics of European two-axle Freight Wagons*. PhD thesis, IMM, The Technical University of Denmark, 2006.
- M. Hoffmann, D. E. Petersen, and H. True. *On the Dynamics of a Railway Freight Wagon with UIC Standard Suspension*, pages 153–160. Proc. 7th Conference on Dynamical Systems - Theory and Applications, Dec. 8–11, 2003. Technical University of Łódź, Łódź, Poland, 2003.
- D.W. Jordan and P. Smith. *Nonlinear Ordinary Differential Equations*. Clarendon Press, Oxford, UK, 1988.
- J. J. Kalker. *Wheel-rail rolling contact theory*, pages 243–261. Mechanics and Fatigue in Wheel/Rail Contact, Proceedings of the Third International Conference on Contact Mechanics and Wear of Rail/Wheel Systems. Elsevier, Amsterdam-New York-Oxford-Tokyo, 1991.
- W. Klingel. Über den Lauf der Eisenbahnwagen auf gerader Bahn. *Organ für die Fortschritte des Eisenbahnwesens in technischer Beziehung, Neue Folge*, XX(4), 1883.

- K. Knothe, A. Groß-Thebing, and H. Le The. *Extension of the Classical Theory of Contact Mechanics to Non-Steady-State and Non-Hertzian Problems*, pages 305–332. Advanced Railway Vehicle System Dynamics. Wydawnictwa Naukowo-Techniczne, Warsaw, Poland, 1991.
- J.D. Lambert. *Computational Methods in Ordinary Differential Equations*. John Wiley & Sons, New York, 1972.
- R.I. Leine and H. Nijmeijer. *Dynamics and Bifurcations of Non-smooth Systems*. Springer Verlag, Berlin, 2004.
- M. Marek and I. Schreiber. *Chaotic Behaviour of Deterministic Dissipative Systems*. Cambridge University Press, Cambridge, UK, 1991.
- F.C. Moon. *Chaotic and Fractal Dynamics*. John Wiley & Sons, New York, 1992.
- A.H. Nayfeh and B. Balachandran. *Applied Nonlinear Dynamics*. John Wiley & Sons, New York, 1995.
- W.A. Nørsett and P.G. Thomsen. Imbedded SDIRK-methods of basic order three. *BIT*, 24:634–646, 1984.
- G. Öderlind. The automatic control of numerical integration. *CWI Quarterly*, 112(1): 55–74, 1998.
- E. Ott. *Chaos in Dynamical Systems*. Cambridge University Press, Cambridge, UK, 1993.
- J-P. Pascal and G. Sauvage. *New Method for reducing the multicontact Wheel/Rail problem to one equivalent rigid contact patch*, pages 475–489. Proc. 12th IAVSD Symposium on Vehicle System Dynamics, The Dynamics of Vehicles on Roads and Tracks. Swets & Zeitlinger, Lisse, 1992.
- D.E. Petersen and M. Hoffmann. Curving dynamics of railway vehicles. Technical report, IMM, The Technical University of Denmark, DK-2800 Kgs.Lyngby, Denmark, 2002.
- J. Piotrowski and H. Chollet. Wheel-rail contact models for vehicle system dynamics including multi-point contact. *Vehicle System Dynamics*, 43(6–7):455–483, 2005.
- J. Piotrowski and W. Kik. *On Calculation of Creep Forces generated within weakly curved area of Contact of Wheel and Rail*, pages 98–108. Bahn-Bau '98, October 28–30, 1998. Verband Deutscher Eisenbahningenieure e. V., Frankfurt/Main, Germany, 1998.
- O. Polach. *A Fast Wheel-Rail Forces Calculation Computer Code*, pages 728–739. Proc. 16th IAVSD Symposium on Vehicle System Dynamics, The Dynamics of Vehicles on Roads and Tracks. Swets & Zeitlinger, Lisse, 2000.
- X. Quost, M. Sebes, A. Eddhahak, J-B. Ayasse, H. Chollet, P-E. Gautier, and F. Thouverez. Assessment of a semi-hertzian method for determination of wheel-rail contact patch. *Vehicle System Dynamics*, 44(10):789–814, 2006.
- G. Rill. A modified implicit EULER algorithm for solving vehicle dynamic equations. *Multibody System Dynamics*, 15(2):1–24, 2006.
- G. Sauvage and J-P. Pascal. Solution of the multiple wheel and rail contact dynamic problem. *Vehicle System Dynamics*, 19:257–272, 1990.
- Z.Y. Shen, J.K. Hedrick, and J.A. Elkins. *A Comparison of alternative Creep-Force models for Rail Vehicle Dynamical Analysis*, pages 591–605. Proc. 8th IAVSD Symposium on Vehicle System Dynamics, The Dynamics of Vehicles on Roads and on Tracks, Cambridge, MA, August 14–19, 1983. Swets & Zeitlinger, Lisse, 1984.

- E.C. Slivsgaard. *On the Interaction between Wheels and Rails in Railway Dynamics*. PhD thesis, IMM, The Technical University of Denmark, 1995.
- E.C. Slivsgaard and H. True. *Chaos in Railway-Vehicle Dynamics*, pages 183–192. Non-linearity and Chaos in Engineering Dynamics. John Wiley & Sons Ltd., Chichester, 1994.
- J.M.T. Thompson and H.B. Stewart. *Nonlinear Dynamics and Chaos*. John Wiley & Sons, New York, 2002.
- H. True. *Does a Critical Speed for Railroad Vehicles exist?*, pages 125–131. RTD-Vol. 7, Proc. of the 1994 ASME/IEEE Joint Railroad Conference, Chicago Ill., March 22–24, 1994. American Society of Mechanical Engineers, United Engineering Center, 345 East 47th Street, New York 10017, USA, 1994a.
- H. True. *On a new Phenomenon in Bifurcations of Periodic Orbits*, pages 327–331. Dynamics, Bifurcation and Symmetry, New Trends and New Tools. NATO ASI Series. Kluwer Academic Publishers, P.O. Box 322, NL-3300 AH Dordrecht, The Netherlands, 1994b.
- H. True. On the theory of nonlinear dynamics and its applications in vehicle systems dynamics. *Vehicle System Dynamics*, 31(5–6):393–421, 1999.
- H. True. Nichtlineare Schienenfahrzeugdynamik, neue Grundlagen, Methoden und Ergebnisse. *ZEVrail Glasers Annalen*, 128(11–12):526–537, 2004.
- H. True and R. Asmund. The dynamics of a railway freight wagon wheelset with dry friction damping. *Vehicle System Dynamics*, 38:149–163, 2002.
- H. True, T.G. Hansen, and H. Lundell. *On the Quasi-Stationary Curving Dynamics of a Railroad Truck*, pages 131–138. RTD-Vol. 29, ASME/IEEE Joint Rail Conference, Pueblo, Col., March 16–18, 2005. American Society of Mechanical Engineers, ASME, 3 Park Avenue, New York, New York 10016, USA, 2005.
- H. True and L. Trzepacz. *The Dynamics of a Railway Freight Wagon Wheelset with Dry Friction Damping in the Suspension*, pages 93–103. Proc. 8th Miniconf. on Vehicle System Dynamics, Identification and Anomalies, Budapest, Nov. 11–13, 2002. Technical University of Budapest, Budapest, Hungary, 2003.
- S. Wiggins. *Global Bifurcations and Chaos*. Springer Verlag, Berlin, 1988.
- S. Wiggins. *Introduction to Applied Nonlinear Dynamical Systems and Chaos*. Springer Verlag, Berlin, 1990.
- F. Xia and H. True. *On the Dynamics of the Three-Piece-Freight Truck*, pages 149–159. RTD-Vol. 25, IEEE/ASME Joint Rail Conference, Chicago Ill., April 22–24, 2003. American Society of Mechanical Engineers, United Engineering Center, 345 East 47th Street, New York 10017, USA, 2003.
- K. Zboinski and M. Dusza. *Analysis and method of the analysis of non-linear lateral stability of railway vehicles in curved track*, pages 222–231. Proc. 18th IAVSD Symposium on Vehicle System Dynamics, The Dynamics of Vehicles on Roads and Tracks. Taylor and Francis, London, UK, 2004.

# Tyre Models, Propulsion and Handling of Road Vehicles

Peter Lugner

Institute of Mechanics and Mechatronics, University of Technology Vienna

**Abstract** As one of the basic essential components for the vehicle dynamics the tyre and its different kinds of modelling are presented first. To characterize the overall dynamic behaviour of a vehicle (passenger car) different simplified vehicle models combined with linearized tyre models are used. Thereby the mathematical formulations for the longitudinal dynamics provide the information for the driving performance and the braking capability of the car while a separate linearized 2-wheel model is introduced to explain and interpret the lateral dynamics and handling characteristics. The application of advanced vehicle and tyre models is discussed to show the possibilities of investigations for nonlinear three-dimensional dynamics.

## 1 Introduction

Since the propulsion properties and the handling of a car are strongly determined by the force transfer between tyre and road it is essential to know also at least the main characteristics of the tyre. For the simulation of e.g. the cornering of a car or the operation of an ABS-system the mathematical-mechanical description of these characteristics, a tyre model is necessary.

Investigations with respect to the properties and the dynamic behaviour of a car can be done using relative simple models – for an easier understanding and interpretation of basic characteristics – or by complex multi body system (MBS) software, that can include detailed design features and material properties. This later kind of models needs the knowledge of an increased number of parameters or property descriptions to get reliable simulation results. Also the tyre models must be described in such detail that they fit to the considered frequency range of the whole system.

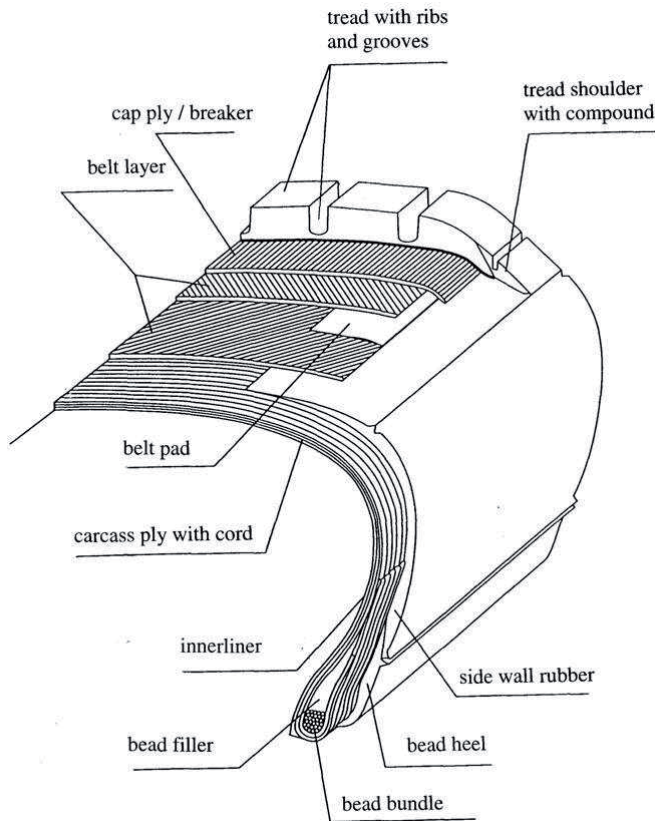
The main goal of the following presentation is to show the overall dynamic behaviour of a passenger car emphasizing the basic components and how they contribute to a good driving and handling behaviour. For completion, some remarks on complex models provide an outlook on further, more detailed investigations.

Some of the notations and the terminology used with vehicle dynamics and tyre properties are standardized in SAE J670e, see e.g. [1] and also provide a good first information of the phenomena of regular driving.



## 2 Tyre models

Since the structure of a tyre, see as an example in Fig. 1 a passenger car tyre, is very complex, all mathematical-mechanical descriptions are directly or indirectly based on measurements.



**Figure 1.** Structural components of a radial passenger car tyre with belt (belt layers plus padding) and carcass (all exclusiv tread and cap ply)

The force transfer between tyre and road surface is the result of the pressure distribution in the contact patch and the tyre slip. The lateral component of the slip is described by the tyre side slip angle, see Fig. 2 and 3. In these figures the kinematics, including the slip values and the essential forces acting on the tyre and the aligning moment  $M_s$ , with respect to the fictive contact point  $A$  are demonstrated.

As can be seen by Fig. 3, the over all sliding velocity  $v_g$  – generally locally there exist sticking and sliding areas in the contact patch, see also Fig. 8 – is used to define the slip

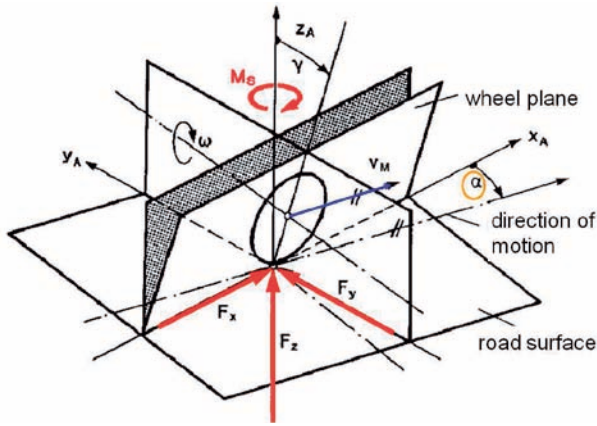


Figure 2. Kinematics and forces from the surface to the wheel

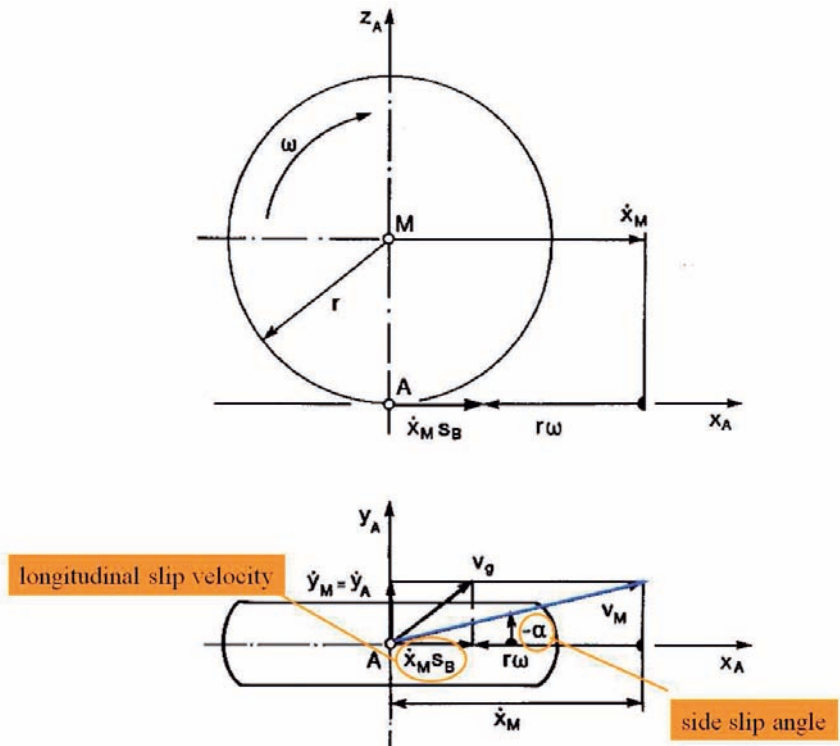


Figure 3. Velocities of the tyre/wheel for camber angle  $\gamma = 0$

values in the following way:

$$\begin{aligned}
 v_g & \text{resultant sliding velocity} \\
 \dot{y}_M & \text{component of the sliding velocity} \\
 \dot{x}_M - r\omega = s_B \dot{x}_M & \text{component of the sliding velocity} \\
 s_B = \frac{\dot{x}_M - r\omega}{\dot{x}_M} = -s_x & \text{braking slip (longitudinal slip with } \dot{x}_M \geq r\omega) \quad (2.1) \\
 s_T = \frac{r\omega - \dot{x}_M}{r\omega} & \text{driving slip (longitudinal slip with } \dot{x}_M \leq r\omega); \\
 & \text{for small values } s_T \cong s_x \\
 s_y = -\frac{\dot{y}_M}{\dot{x}_M} = \tan \alpha & \text{lateral slip and side slip angle } \alpha \text{ of the wheel.} \quad (2.2)
 \end{aligned}$$

Measurements of the tyre properties are mostly done on roller test rigs, see [2], but also road measurements are available. Two typical measurements for two different passenger car tyres are shown in Fig. 4 and 5. These quasi steady state measurements are characterized by a relatively strong increase of the lateral force  $F_y$  with the side slip angle  $\alpha$ , with  $F_y$  reaching its maximum at about  $\alpha = 10^\circ$ - $12^\circ$ . With higher values of  $\alpha$  it slightly decreases till it reaches its value for complete sliding. The behaviour of the longitudinal force  $F_x$ , see Fig. 7 with simulated results, with respect to the longitudinal slip is very similar though  $F_x$  reaches its maximum at lower longitudinal slip values of about  $|s_x| = 0, 1$ .

For a numerical simulation of the vehicle behaviour the tyre characteristics need to be modeled in such a way that it fits to the frequency range and complexity of the vehicle model.

So for low frequency handling and driving manoeuvres a linear approximation is the first choice, see Fig. 6, for the lateral force  $F_y$  versus side slip angle. The mathematical description uses the cornering shiftiness  $C_\alpha$ . And the longitudinal force transfer is described in a similar way

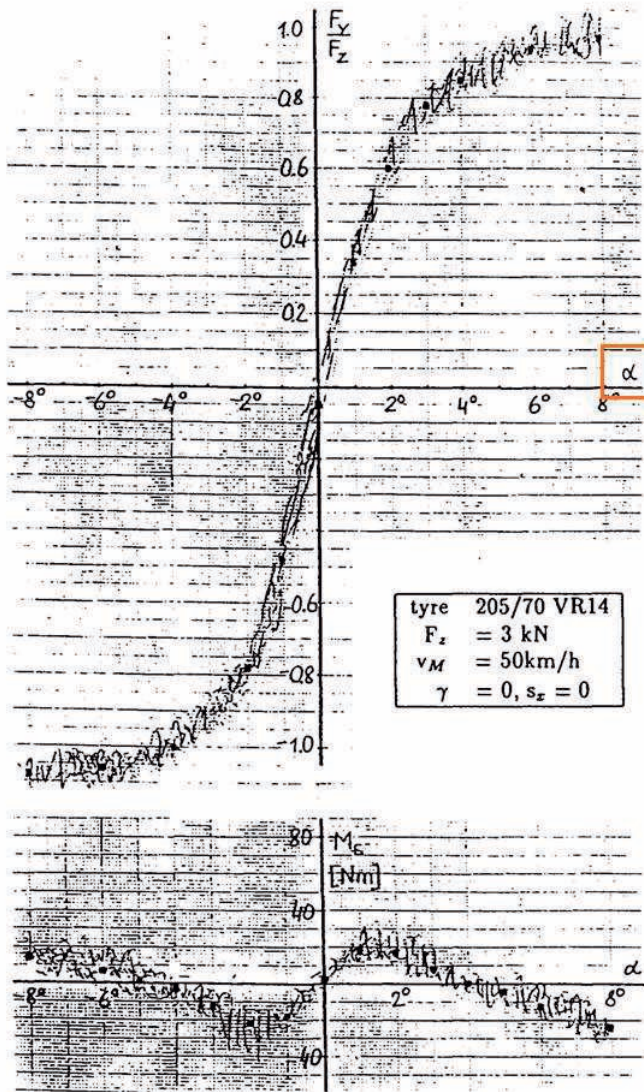
$$F_y = C_\alpha \alpha \quad (2.3)$$

$$F_x = C_x s_x. \quad (2.4)$$

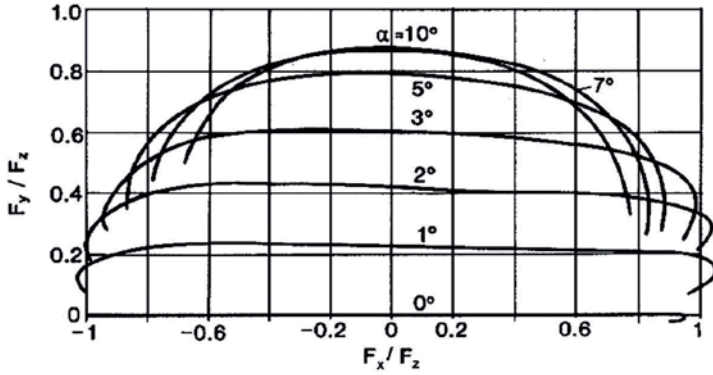
It should be noticed that this kind of approximation does not include the physical limitation of the tyre forces nor the mutual influence of lateral and longitudinal forces like obvious by Fig. 5.

Other more comprehensive approximations of the measured tyre behaviour can be found in [3], [4], [5]. Based on semi empirical considerations the full nonlinearities are taken into account and the whole range of possible working conditions of the wheel/tyre-road contact can be exploited. An example based on the normalization of the measurements with respect to maximum and corresponding slip values, [5], is shown in Fig. 7. The also stated principal functions for the steady state values of lateral force  $F_y^s$ , longitudinal force  $F_x^s$  and aligning torque  $M_s^s$  want to indicate the main variables. Since these kinds of approximations do not include tyre dynamics by the actual force values  $F_i$  the principal transient phenomena can be approximated or added by

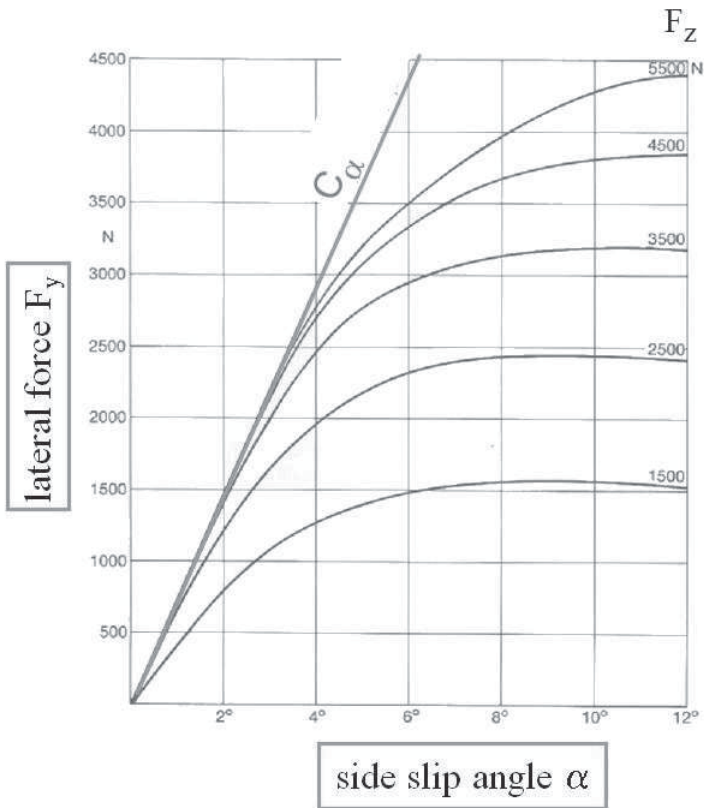
$$\dot{F}_i = \frac{v_M}{l_i} (F_i^s - F_i) \quad i = x, y \quad (2.5)$$



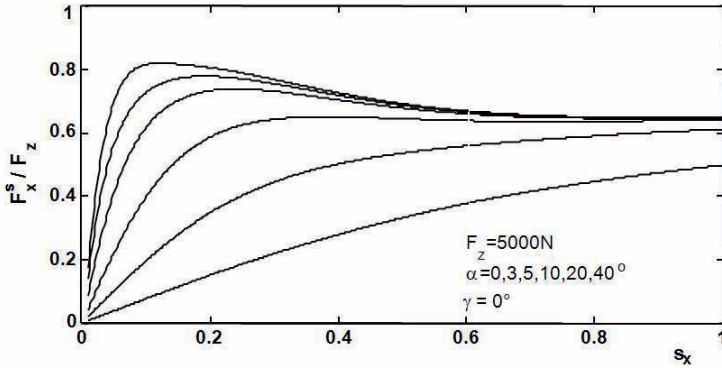
**Figure 4.** Steady state lateral force  $F_y$  and self aligning torque  $M_s$  from road measurements.



**Figure 5.** Combined force transfer due to side slip angle  $\alpha$  and longitudinal force  $F_x$  for constant normal force  $F_z$ , test rig measurements.



**Figure 6.** Lateral force  $F_y$  and linear approximation by the cornering stiffness  $C_\alpha = \left(\frac{\partial F_y}{\partial \alpha}\right)_{\alpha=0}$ .



$$F_y^s = F_y^s(\alpha, s_x, \gamma, F_z), \quad F_x^s = F_x^s(\alpha, s_x, \gamma, F_z), \quad M_s^s = M_s^s(\alpha, \gamma, F_z, F_y^s, F_x^s)$$

**Figure 7.** Simulation of the steady state longitudinal force  $F_x^s$  as function of side slip angle  $\alpha$  and longitudinal slip  $s_x$  and principal functional description of the tyre forces and the aligning torque  $M_s^s$ .

with transition length  $l_i$ .

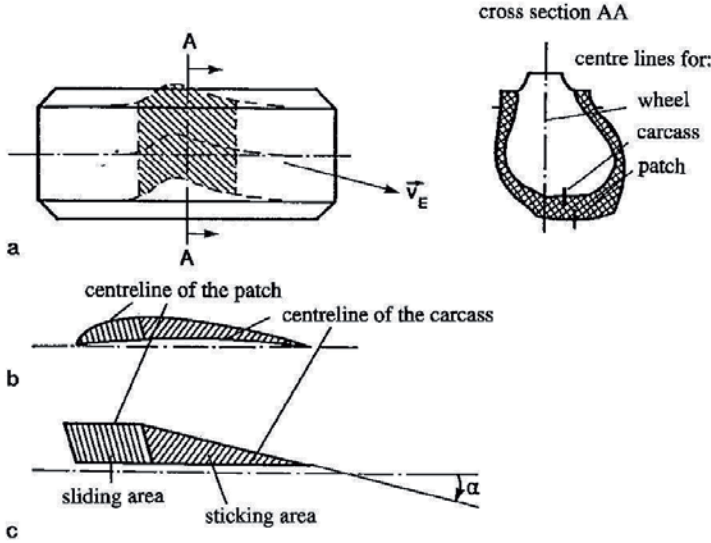
An approach to include the tyre deformations in the kind of modeling was introduced by the HSRI-model, see Fig. 8, [6], [7], [8].

Based on the contribution of sliding and sticking areas shown in Fig. 8 a mathematical formulation can be derived using a relative small number of tyre parameters, Fig. 9.

In this HSRI tyre model the friction value, which is responsible for the distribution of sliding and sticking areas, is a function of the sliding velocity and can be adapted to the surface conditions by the parameters  $f_o$ ,  $k_R$ . The influence of the longitudinal force  $F_x$  on the self aligning torque  $M_s$  contains the lateral tyre deflection by  $F_y/s_y$ . The tyre load  $F_{z,stat}$  represents a nominal value and by  $F_z/F_{z,stat}$  an adaption with respect to the actual vertical load can be accomplished. A transient tyre behaviour could be added similar to equation 2.5. Fig. 10 shows a typical lateral force behaviour of a passenger car tyre calculated by the HSRI model.

For more detailed investigations, e.g. running over a cleat and higher frequency regions up to about 60 Hz the tyre must be represented by a more complex model. As an example, the structure of the tyre model SWIFT, [3], includes a spring-damper compliance between a rigid belt ring and the wheel rim and an additional residual spring between belt and contact patch, see Fig. 11. To be able to detect road surface steps and calculate the corresponding tyre forces especially  $F_x$ ,  $F_z$ , a sensing scheme with cams, that substitute the contours of the deflected tyre, is introduced to determine an effective road plane and approximately is curvature.

Spring and damper nonlinearities are taken into account, e.g. in vertical direction. The belt to rim vibration eigenfrequencies take into account influences of rolling speed and vertical load  $F_z$ . The transient behaviour is described by first order differential equations for the transient longitudinal slip  $s'_x$ , side slip angle  $\alpha'$  and a transient  $\alpha'_t$  for



**Figure 8.** Deformations of tyre and in the tyre patch with velocity  $v_E$  at entrance (side slip angle  $\alpha$ ).

a ... deformation of tyre

b ... simplification of patch area

c ... HSRI modeling

the pneumatic trail  $t_p$ .  $\sigma_i$  are transition lengths and  $v_{sci}$  the components of the sliding velocity of the contact patch (see Fig. 11):

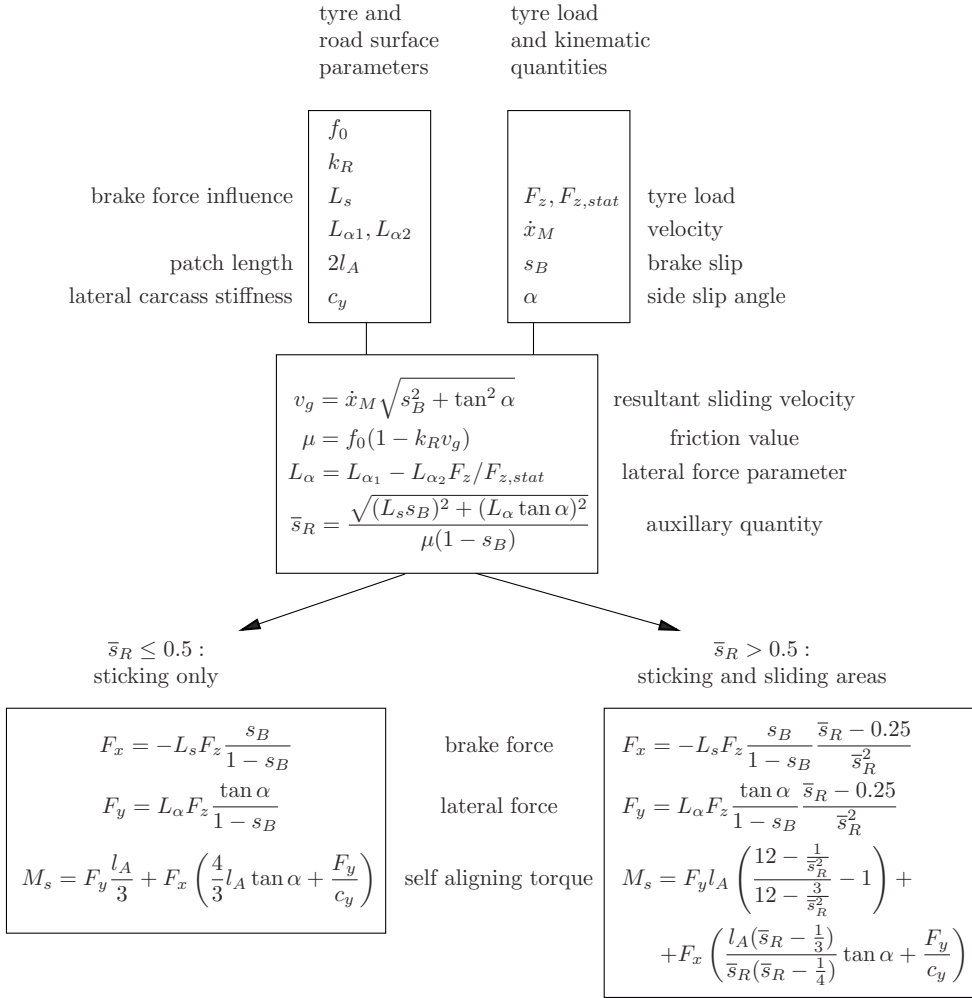
$$\begin{aligned}
 \sigma_c \frac{ds'_x}{dt} + |\dot{x}_M| s'_x &= -v_{scx} \\
 \sigma_c \frac{d\alpha'}{dt} + |\dot{x}_M| \alpha' &= -v_{scy} - |\dot{x}_M| \psi_{st} \\
 \sigma_t \frac{d\alpha'_t}{dt} + |\dot{x}_M| \alpha'_t &= |\dot{x}_M| \alpha'.
 \end{aligned} \tag{2.6}$$

The tyre forces and the self aligning torque  $M_s$  are principally described by nonlinear functions, [10]

$$\begin{aligned}
 F_x &= F_x(s'_x, \alpha', F_z, \dots) \\
 F_y &= F_y(s'_x, \alpha', F_z, \dots) \\
 M_s &= -t_p(s'_x, \alpha'_t, \dots) F_y + M_{zr}(s'_x, \alpha', \dots) + (s + y_{st}) F_x + C_{\Delta M}(\alpha' - \alpha'_t).
 \end{aligned}$$

Besides the component with the pneumatic trail  $t_p$  the aligning torque  $M_s$  includes a moment  $M_{zr}$  caused by a spin slip, the part of the longitudinal force  $F_x$  times lateral deflection and a possible residual component  $C_{\Delta M}$ .

In Fig. 12 a comparison of a SWIFT simulation and the measurement of a straight running tyre with  $\alpha = 0$  is shown. A sequence of 4 different surface obstacles is encoun-





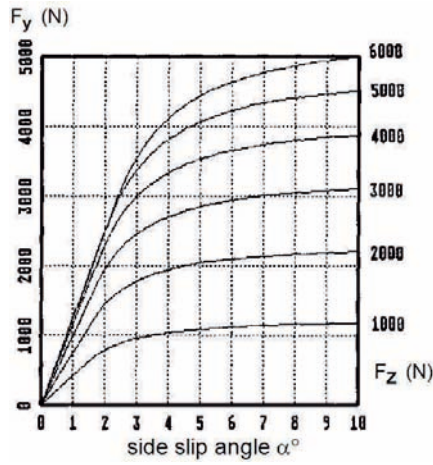


Figure 10. HSRI-model: steady state lateral force  $F_y$  for different tyre loads  $F_z$ .

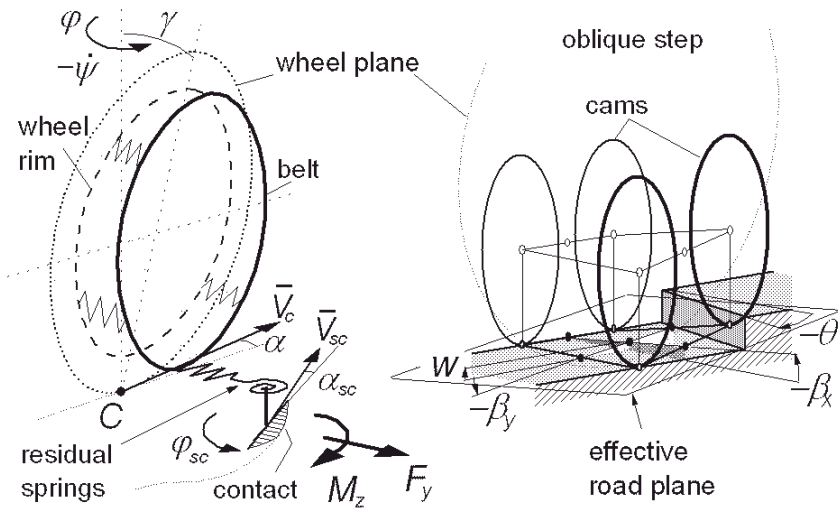
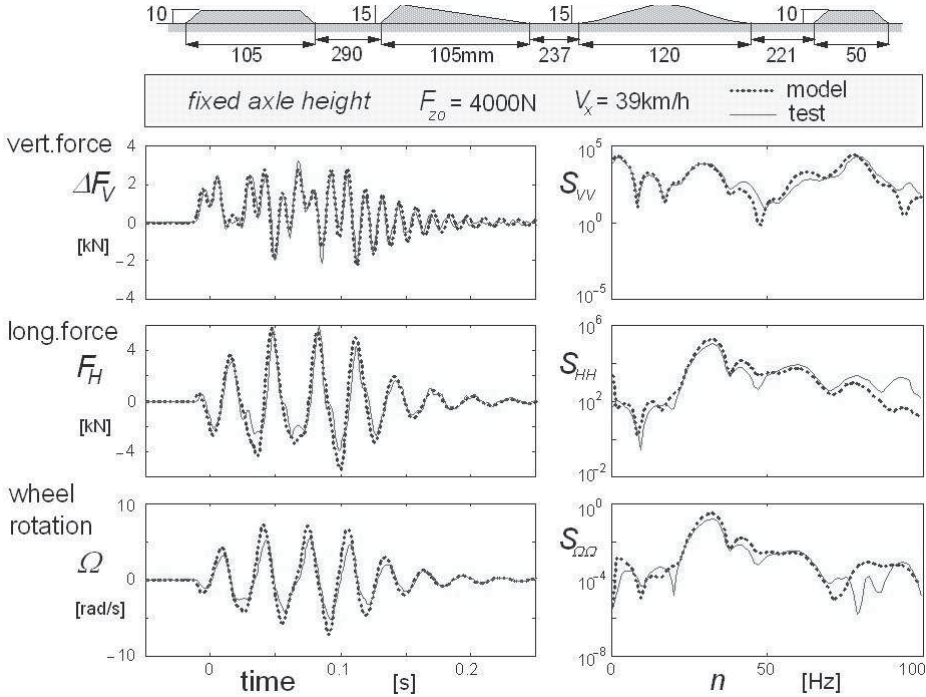


Figure 11. The structure of the 3D tyre model SWIFT.

tered. The differences show a very good agreement for this test, where the axle height of the wheel was fixed. The corresponding power spectral densities  $S$  reveal greater deviations for higher frequencies only.



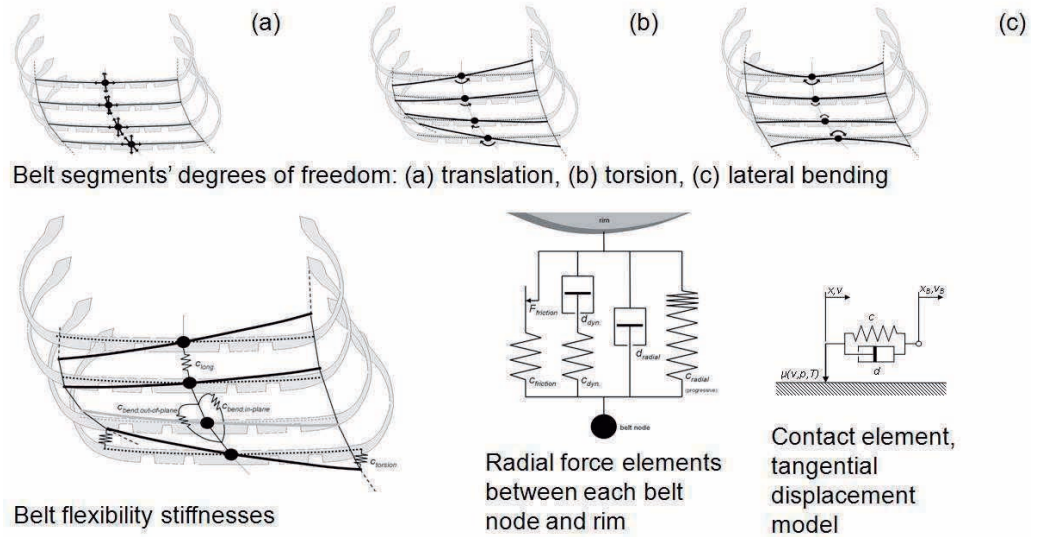
**Figure 12.** SWIFT tyre model: straight running over different obstacles.

The structure of an other 3D-tyre model that moreover includes a flexibility of belt elements is shown in Fig. 13, [9].

The main characteristics of this model are the structural stiffness, damping and inertia properties and also the calculation of the tread/road contact with pressure distribution and local friction. For longitudinal encountered obstacles wave length to 0,05m, including sharp edges, are fully simulated, in lateral direction wave length to about 0,15m. It is possible to take into account temperature and road friction changes directly and also changes of the inflation pressure.

Fig. 14 shows measurements and simulations of the SWIFT as well as of the FTire models when running over a rectangular 20x20mm cleat at two different velocities. Both models demonstrate the ability to reproduce the dynamic effects thought for the SWIFT model the first peaks in the longitudinal force  $F_x$  for the higher velocity are obviously larger then those of the measurements.

As already said for the simulation of the vehicle behaviour, it is very essential to use an appropriate tyre model description. Quite a lot of basic considerations, as shown in the following chapters, can be based on very simple tyre force transfer assumption – sometimes even with neglecting the slip at all – but with increasing vehicle model



**Figure 13.** Structure of the tyre model FTire.

complexity the tyre model structure needs to become also more detailed to be able to describe the higher frequency phenomena. But it needs to be mentioned that with higher tyre model complexity the number of necessary parameters or characteristic functions based on measurements increase and moreover the use of such model with additional degrees of freedom due to the belt-rim movements or flexibility of structural components with a multi-body-system software should be checked by special test procedures to be sure that it is working according to the investigated problem, [10].

### 3 Longitudinal Dynamics

Low frequency vehicle dynamics with linearized models provide the basics for the interpretation of the global vehicle behaviour. For these investigations different (nearly) independent models for longitudinal, lateral and vertical dynamics (ride and comfort) can be used, e.g. [1], [11], [12].

#### 3.1 Vehicle and System Models

For the longitudinal dynamics the phenomena like accelerating and braking will be considered. Thereby the engine performance with fuel consumption will not be included. Furthermore it is assumed that generally the longitudinal slip values are very small or the wheels operate with pure rolling. Possible dynamic effects by the car body roll, heave or pitch are neglected for this basic approach to provide simple but substantial information with respect to the motion of a car along a trajectory.

Fig. 15 shows the model of the passenger car running on a flat slope with inclination. Here and further on a symmetry with respect to the car central  $xz$ -plane is presupposed.

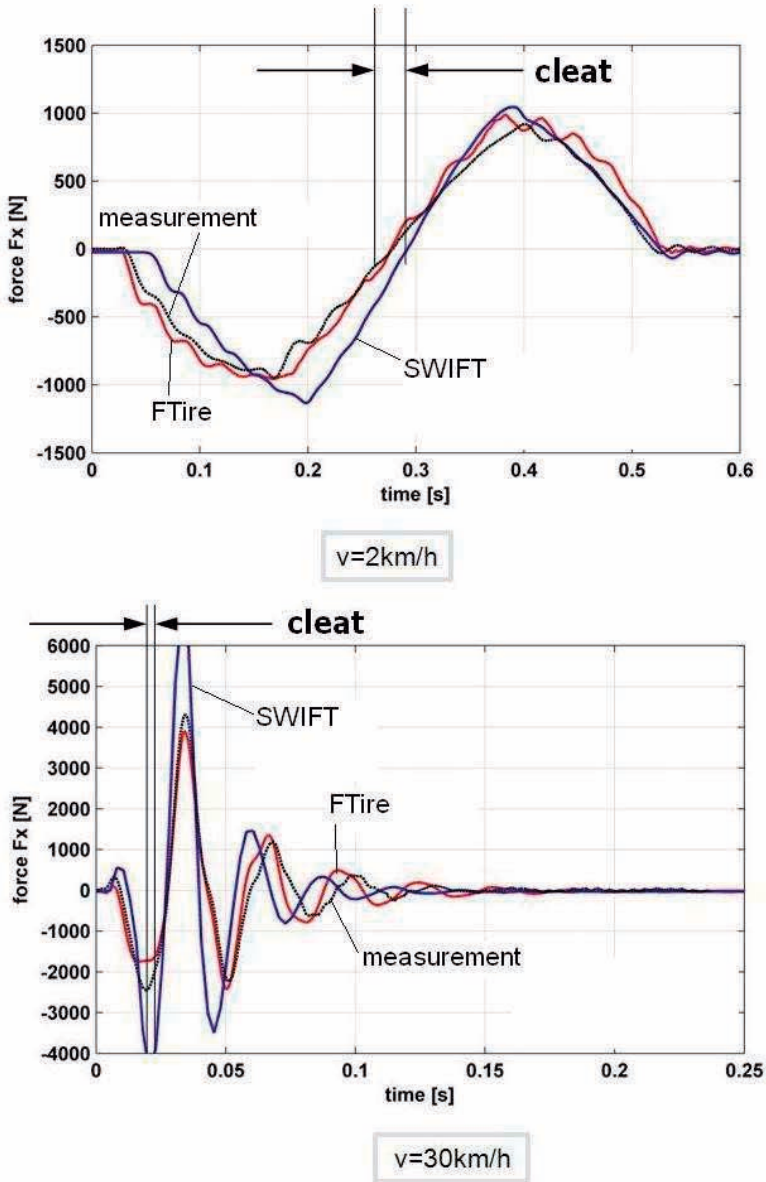
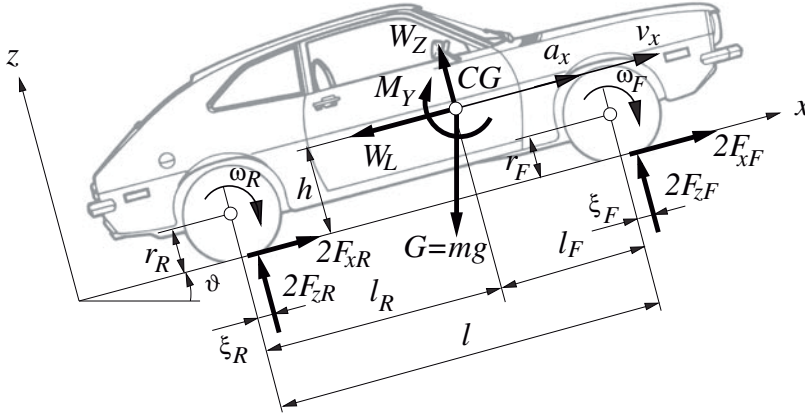


Figure 14. Comparison of measurements with the simulations by SWIFT and FTire for running over a rectangular cleat normal to the running direction,  $\alpha = 0$ .

The meanings of most symbols are self explaining;  $\xi_F$  and  $\xi_R$  are the consequences of the rolling resistances of front and rear tyres, see Fig. 19. As an example in Fig. 16 a



**Figure 15.** Plane longitudinal vehicle model.

rear wheel is extracted from the vehicle model and the forces and torques acting at this wheel, also those from the car to the wheel, are pictured.

The equations of motion for the vehicle derived by Fig. 15, with the overall mass  $m$  of the vehicle, are

$$ma_x = 2F_{xF} + 2F_{xR} - W_L - G \sin \vartheta \quad (3.1)$$

$$0 = 2F_{zF} + 2F_{zR} + W_Z - G \cos \vartheta \quad (3.2)$$

$$2I_F \dot{\omega}_F + 2I_R \dot{\omega}_R = 2F_{zR}(l_R - \xi_R) - 2F_{zF}(l_F + \xi_F) - 2(F_{xF} + F_{xR})h + M_Y \quad (3.3)$$

and the Euler equation for the rear wheel (by Fig. 16), with the moment of inertia  $I_R$  with respect to the wheel axis:

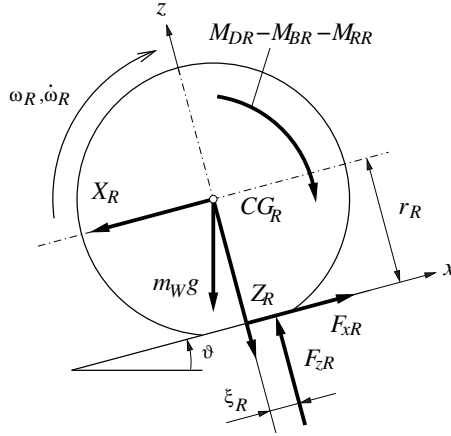
$$I_R \dot{\omega}_R = M_{DR} - M_{BR} - M_{FR} - F_{zR} \xi_R - F_{xR} r_R. \quad (3.4)$$

With the neglect of a longitudinal slip and the assumption of negligible differences in the rolling radius  $\bar{r}$  and the deflection radii

$$\bar{r} = r_F = r_R = r \quad (3.5)$$

the simplified rolling conditions for all wheels become ( $v_x \cong v$ ):

$$\begin{aligned} \omega &= \omega_F = \omega_R \\ r\omega &= v, \quad r\dot{\omega} = a_x. \end{aligned} \quad (3.6)$$



- $M_{DR}$ ... driving torque
- $M_{BR}$ ... braking torque
- $M_{FR}$ ... friction torque of the wheel mounting
- $X_R, Z_R$ ... reaction force from the vehicle

**Figure 16.** Model of a rear wheel corresponding to the car model of figure 15.

Inserting from (3.4) for the rear wheel longitudinal force  $F_{xR}$  and the analogue expression for  $F_{xF}$  in (3.1) provides

$$\begin{aligned}
 ma_x + 2I_R \frac{\dot{\omega}}{r} + 2I_F \frac{\dot{\omega}}{r} &= \frac{1}{r} [(2M_{DR} + 2M_{DF}) - (2M_{BR} + 2M_{BF})] \\
 -2 \left[ \left( \frac{M_{FR}}{r} + \frac{M_{FF}}{r} \right) + (f_R F_{zR} + f_R F_{zF}) \right] &- W_L - mg \sin \vartheta
 \end{aligned} \tag{3.7}$$

where the overall rolling resistance coefficient, (see chapter 3.2) is used:

$$f_R = \frac{\xi}{r}.$$

With the rolling condition (3.6) the equation (3.7) for the longitudinal acceleration can be written in a compact form:

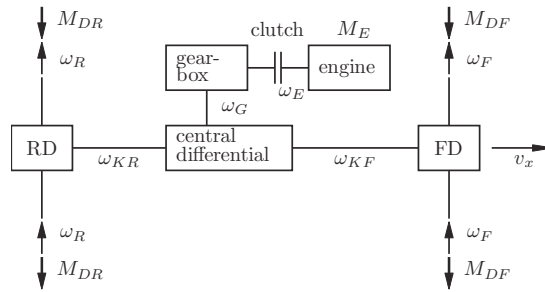
$$\left( m + \frac{2I_R}{r^2} + \frac{2I_F}{r^2} \right) a_x = \frac{M_D}{r} - \frac{M_B}{r} - W_{ges} \tag{3.8}$$

with the substitutes

$$\begin{aligned}
 W_{ges} &= \frac{M_F}{r} + W_R + W_L + W_G \\
 W_R &= f_R(2F_{zR} + 2F_{zF}) = f_R(G \cos \vartheta - W_Z) \\
 W_G &= mg \sin \vartheta \\
 M_F &= 2(M_{FR} + M_{FF}) \\
 M_D &= 2M_{DR} + 2M_{DF} \\
 M_B &= 2M_{BR} + 2M_{BF}.
 \end{aligned}$$

The components of the overall resistance  $W_{ges}$  comprise the overall rolling resistance  $W_R$ , friction moments  $M_F$  of the wheel bearings (which generally can be neglected), the aerodynamic resistance  $W_L$  and the down hill gravity component  $W_G$ . With the drive moments  $M_D$  and the braking moment  $M_B$  the possible input actions by the driver via engine and brakes can be described.

For providing the drive moments  $M_{DF}, M_{DR}$  for the front and the rear axle the drive train is considered in a simplified form, see Fig. 17. A general 4WD moment and angular velocity distribution is presented in [19].



**Figure 17.** Drive train for 4WD with equal running conditions for left and right side.

Thereby  $\omega_E$  is the angular velocity of the engine and  $\omega_{KF}, \omega_{KR}$  those of the front and rear cardan shafts of the central differential. With distribution factors  $\nu_F, \nu_R$  now different kinds of drive concepts can be considered. With the use of the overall drive torque  $M_D$ , see (3.8) it can be written

$$\begin{aligned}
 2M_{DF} = M_D \nu_F, \quad 2M_{DR} &= M_D \nu_R \\
 \text{with } \nu_F + \nu_R &= 1
 \end{aligned}$$

$$\begin{aligned}
 \text{for rear wheel drive} &: \quad \nu_R = 1 \\
 \text{for front wheel drive} &: \quad \nu_F = 1 \\
 \text{for 4WD with equal distribution} &: \quad \nu_R = \nu_F = 0,5.
 \end{aligned} \tag{3.9}$$

The corresponding kinematics are

$$\begin{aligned}\omega_E &= \omega_G N_{Gn}, \\ \omega_G &= \nu_R \omega_{KR} + \nu_F \omega_{KF} \\ \omega_{KR} &= N_D \omega_R, \quad \omega_{KF} = N_D \omega_F\end{aligned}\tag{3.10}$$

with the transmission ratio  $N_D$  for the axle differentials and the ration  $N_{Gn}$  for the chosen gear (by the gear box).

Considering now the torque transmission from the engine tho the wheels including the moments of inertia  $I_C, I_{DF}, I_{DR}$  of parts of the gears and central differentials, and also parts connected to the cardan shafts and the substitutive inertia  $\Theta$  of the engine, provides the equation (for 4WD):

$$\begin{aligned}M_D &= (2M_{DF} + 2M_{DR}) = \\ &\eta M_E(\omega_E) \cdot N - r \left[ \frac{\Theta_E N^2}{r^2} + \frac{(I_C + I_{DR} + I_{DF}) N_{Gn}^2}{r^2} \right] a_x\end{aligned}\tag{3.11}$$

with

$$N = N_{Gn} N_D.$$

The coefficient  $\eta < 1$  wants to take into account the energy losses in the transmission system. The engine torque  $M_E$  will be a function of the engine speed  $\omega_E$ .

Now combining (3.8) with (3.11) leads to the essential equation for the longitudinal acceleration and dynamical behaviour of the car:

$$m \lambda a_x = \frac{\eta M_E \cdot N}{r} - \frac{M_B}{r} - W_{ges}\tag{3.12}$$

with the mass factor  $\lambda$  and the reduced mass  $m_r$  of the rotational parts of the drive train and wheels

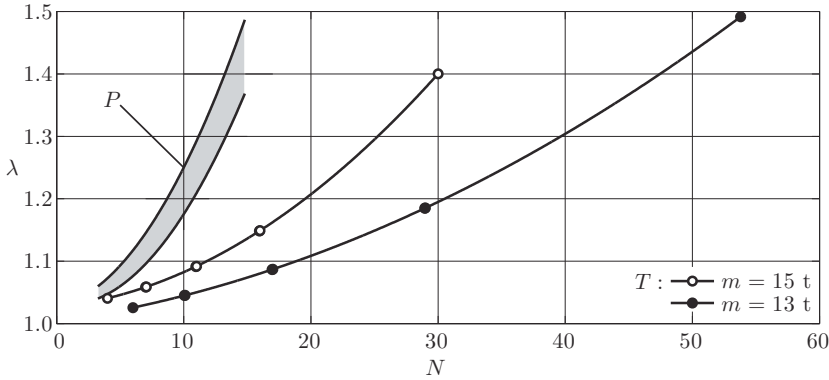
$$\begin{aligned}\lambda &= \frac{m + m_r}{m} \\ m_r &= \frac{1}{r^2} [\Theta_E N^2 + (I_C + I_{RD} + I_{FD}) N_{Gn}^2 + (2I_R + 2I_F)].\end{aligned}\tag{3.13}$$

Examples for  $\lambda$  for special types of vehicles are shown in Fig. 18, [12].

For the braking performance and generally for the force transfer at the tyres the normal forces, see Fig. 15, 16, need to be known. Using (3.1) till (3.3) and  $\xi_F = \xi_R = \xi$ , two equations can be established for the determination of the normal forces  $F_{zF}, F_{zR}$ :

$$\begin{aligned}2F_{zF} + 2F_{zR} &= G \cos \vartheta - W_z \\ -2l_F F_{zF} + 2l_R F_{zR} &= \left[ \left( \frac{2I_F}{r} + \frac{2I_R}{r} \right) a_x + (m a_x + W_L + G \sin \vartheta) h \right. \\ &\quad \left. - M_Y + \xi (G \cos \vartheta - W_Z) \right]\end{aligned}\tag{3.14}$$





**Figure 18.** Typical mass factor  $\lambda$  as function of overall transmission ratio  $N$  for passenger cars (P) and trucks (T).

For principal considerations it is useful to neglect small terms and linearize trigonometric functions where possible. So with  $M_Y \cong 0$ ,  $W_Z \cong 0$ ,  $\xi(G \cos \vartheta) \ll (ma_x + W_L + G \sin \vartheta)h$  and  $(2I_F + 2I_R)/r \ll mh$  the simplified expressions for the normal forces become

$$\begin{aligned} \frac{F_{zF}}{G} &= \frac{l_R}{l} \cos \vartheta - a^* \frac{h}{l} \\ \frac{F_{zR}}{G} &= \frac{l_F}{l} \cos \vartheta + a^* \frac{h}{l} \end{aligned} \quad (3.15)$$

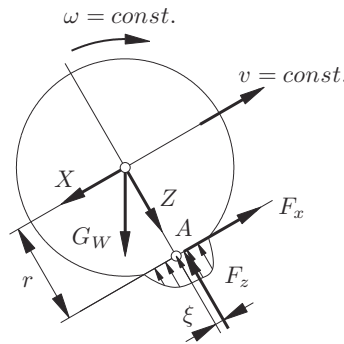
$$\text{with } a^* = \left( \frac{a_x}{g} + \sin \vartheta + \frac{W_L}{mg} \right)$$

### 3.2 Driving resistances

Before describing the rolling resistance and aerodynamic forces as the main components of the driving resistance, a short remark should explain some other, sometimes used approaches. In some literature terms or definitions of "acceleration resistance" or "grade resistance" are used. Besides that the mathematical expressions of these quantities can be negative or positive (which does not compare to "resistance"), the mathematical formulations of e.g. equations (3.8), (3.12) or also (3.15) already include these influences by  $M_D, M_B$  and grade angle  $\vartheta$ . It is therefore recommended to avoid these kind of definitions.

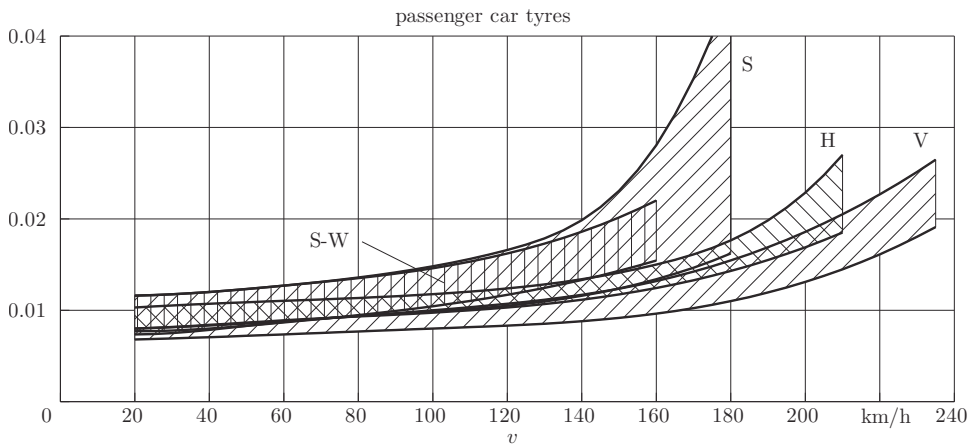
The rolling resistance of the tyre can be explained by the deformation work by forming the patch. As a consequence the resultant normal tyre force  $F_z$  acts a small distance  $\xi$  in front of the fictive contact point  $A$ , see Fig. 19. So if the wheel/tyre is moving with constant velocity  $v$  and constant angular velocity  $\omega$  a force  $F_x$  (and thereby a force  $X$  from the vehicle to the wheel) is necessary to keep this condition.

$$F_x = -\frac{\xi}{r} F_z = -f_R F_z \quad (3.16)$$



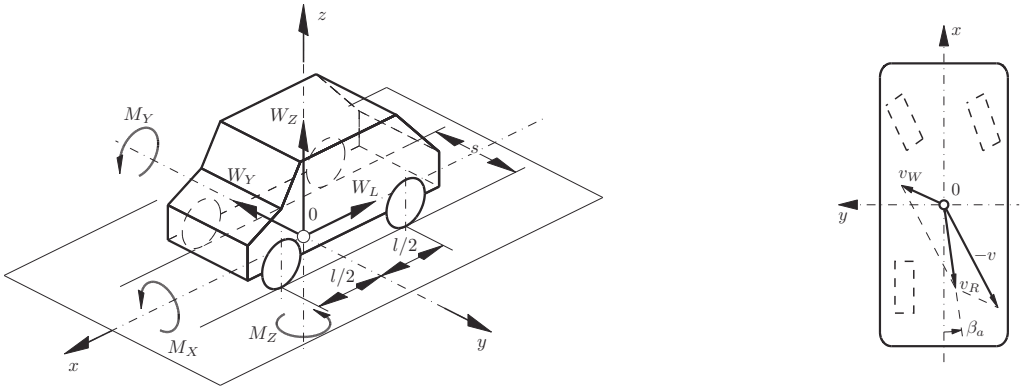
**Figure 19.** Tyre motion with constant velocities.

Typical values for the so defined rolling resistance coefficient  $f_R$  are shown in Fig. 20, [12]. Characteristic is the remarkable increase of  $f_R$  near the admissible speeds of the different types of tyres.



**Figure 20.** Rolling resistance coefficients  $f_R$  for winter tyres S-W, and tyre types S, H, V with different maximum admissible speeds.

The aerodynamic interaction of the vehicle with the air flow around the chassis is characterized by the reduction of this pressure distribution to a reference point 0 – see Fig. 21. The resultant forces and moments are described by the aerodynamic pressure  $q$ , the cross section area  $A$  and appropriate coefficients  $C_i$  and the wheel base  $l$ , e.g. [13], [11].



**Figure 21.** Coordinate system and reference point 0 for the aerodynamic forces and moments and the aerodynamic side slip angle  $\beta_a$  due to ambient wind  $v_w$  and vehicle velocity  $v$ .

$$\begin{aligned}
 W_L &= C_T A \cdot q, & W_Y &= C_Y A \cdot q, & W_Z &= C_Z A \cdot q, \\
 M_X &= C_{M_x} A \cdot ql, & M_Y &= C_{M_y} A \cdot ql, \\
 M_Z &= C_{M_z} A \cdot ql \\
 q &= \frac{\rho v_R^2}{2}
 \end{aligned} \tag{3.17}$$

For the aerodynamic drag  $W_L$  for passenger cars, the tangential coefficient  $C_X = C_T(\beta_a = 0)$  is used, a simplified approach that is based on measurements like shown in Fig. 22. Obviously for trucks and busses especially for larger  $\beta_a$  other approaches have to be used. Often the coefficients  $C_{M_z}$  and  $C_Y$  can be described as linear function of  $\beta_a$ , [12].

Mostly for the simplified vehicle dynamics the drag  $W_L$ , lateral force  $W_Y$  and a aerodynamic yaw moment with respect to the  $CG$  are applied to the vehicle model only.

### 3.3 Driving Performance

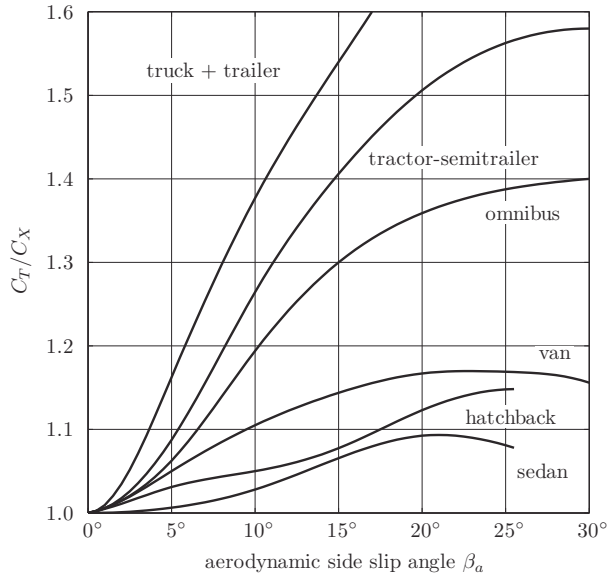
According to equation (3.12) and the knowledge of the driving resistances the capability of the vehicle e.g. to accelerate and what maximum driving speed is possible can be evaluated if the engine characteristics – the torque  $M_E(\omega_E)$  – is known.

Fig. 23 shows typical full throttle power  $P_{max}$  and torque  $M_{E,max}$  characteristics for a Gasoline and Diesel engine, [1].

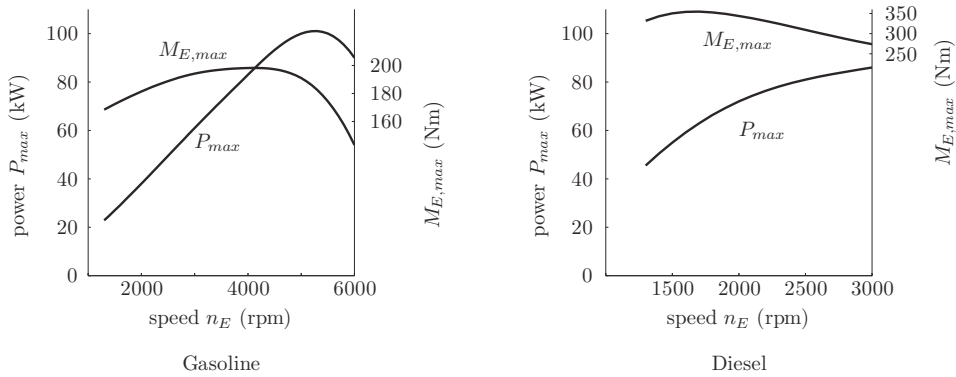
Assuming an linear relationship between throttle position  $\lambda_T$  and the available engine torque  $M_E$  it can be written

$$M_E = (M_{E,max} - M_{E,d})\lambda_T + M_{E,d} \quad 0 \leq \lambda_T \leq 1 \tag{3.18}$$

with the engine drag torque  $M_{E,d} = M_E(\lambda_T = 0)$ . For a better demonstration of the vehicle performance a "driving force  $K_E$ " is introduced, that does not only include the



**Figure 22.** Normalized tangential coefficient  $C_T/C_X$  for different types of vehicles as function of the aerodynamic side slip angle  $\beta_a$ .



**Figure 23.** Examples of full throttle engine characteristics.

engine torque as function of engine speed (and  $\lambda_T$ ) but also the actual gear ratio  $N_{Gn}$ :

$$K_E = \frac{\eta M_E(\omega_E, \lambda_T) \cdot N_{Gn} N_D}{r}, \quad \omega_E = v \cdot \frac{N_{Gn} N_D}{r}. \quad (3.19)$$

Thereby finally the driving performance can be described by

$$m\lambda a_x = K_E(v, \lambda_T, N_{Gn}) - W_{ges}(v, \vartheta). \quad (3.20)$$

With this equation it is assumed, that the slip values for the transfer of the tire forces can be neglected – an assumption that needs to be checked for high accelerations and velocities or grades  $\vartheta$  (by determining  $F_{xi}/F_{zi}$  – see also chapter 3.4).

Derived from (3.20) with engine characteristics similar to figure 23, gasoline engine, the approximation of the engine drag force  $K_{E,d}$  by a linear function and the driving resistances (over all rolling resistance  $f_R G$  and aerodynamic drag  $W_L$ , see chapter 3.2) a drive performance diagram, Fig. 24, can be established. The shown curves for the operation field of the engine are the limitations by the maximum drive force  $K_{E,max}$  and the drag force  $K_{E,d}$  for the 4 different gear ratios  $N_{Gi}$ . Due to the throttle activation the driver can choose any value in between. For level road the overall resistance as function of driving velocity  $v$  is marked with  $W_{ges}, q = 0$ . For grades  $q = 10\%, -10\%$  the curves are parallel to this basic one.

For a better interpretation of this Fig. 24, some special conditions are emphasized. So for  $q = 0$  the vehicle will have a maximum possible speed of  $v_{max}$  as determined by the intersection of  $W_{ges}, q = 0$  and the maximum driving force  $K_{E,max}$  for the 4th gear  $N_{G4}$ . Similarly  $v_1$  characterizes the downhill velocity ( $q = -10\%$ ), 3rd gear, without throttle activation or braking. If the 2nd gear is used at  $v_2$  the vehicle will have an acceleration capacity (for full throttle) according to  $m\lambda a_{x,2}$ . Other driving conditions can be found in a similar way.

### 3.4 Braking Performance

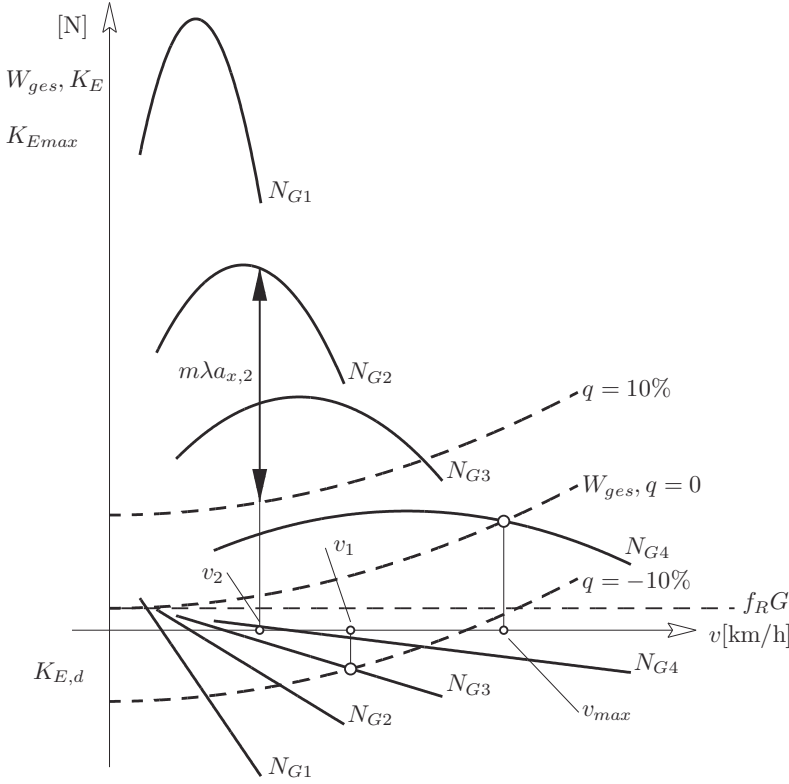
In the vehicle decelerations  $a^*$  also the influences of a grade (small  $\vartheta$  is assumed) and the aerodynamic drag  $W_L$  are considered (corresponding to the forces in the  $CG$ , Fig. 15) – see also (3.14).

$$a^* = \left( \frac{a_x}{g} + \vartheta + \frac{W_L}{mg} \right), \quad \beta = -a^* > 0 \quad (3.21)$$

With  $\beta$  the vehicle deceleration should be characterized with a positive value.

For further investigations it is appropriate to introduce normalized, dimensionless quantities (for horizontal road,  $\vartheta = 0$ )

$$\begin{aligned} \frac{2F_{xi}}{G} &= \bar{F}_{xi}, & \frac{h}{l} &= \bar{h}, \\ \frac{2F_{zi}}{G} &= \bar{F}_{zi}, & \frac{l_i}{l} &= \bar{l}_i, \end{aligned} \quad i = F, R. \quad (3.22)$$



**Figure 24.** Driving performance diagram.

Now for the intended principal considerations the simplified and condensed equations of motion (3.1), (3.15) can be written with

$$-\beta = \bar{F}_{xF} + \bar{F}_{xR} \tag{3.23}$$

$$\bar{F}_{zF} = \bar{l}_R + \beta \bar{h}$$

$$\bar{F}_{zR} = \bar{l}_F - \beta \bar{h}. \tag{3.24}$$

The essential limit for the braking performance is the utilization of the available friction (coefficient  $\mu_{max}$ ). The limit conditions may occur at the front or rear axle or at both axles at the same time.

Starting first with the ideal case that  $\mu_{max}$  is utilized at the same time at both axles. Then the normalized brake forces become

$$\bar{F}_{xR} = -\mu_{max} \bar{F}_{zR}, \quad \bar{F}_{xF} = -\mu_{max} \bar{F}_{zF} \tag{3.25}$$

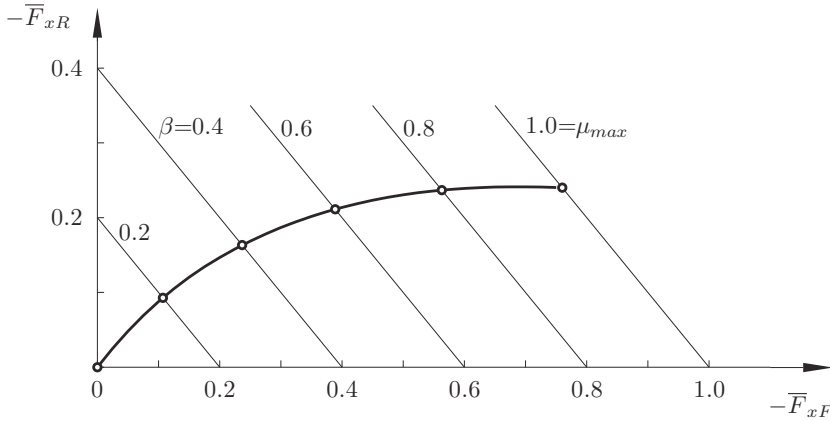
and the normalized deceleration

$$\beta = \beta_{ideal} = \mu_{max}. \tag{3.26}$$

To achieve this condition a brake force balance of

$$\frac{-\bar{F}_{xR}}{-\bar{F}_{xF}} = \frac{\bar{l}_F - \bar{h}\mu_{max}}{\bar{l}_R + \bar{h}\mu_{max}} \quad (3.27)$$

has to be realized. Fig. 25 shows for a specific  $CG$  position the necessary function (3.27), for different  $\mu_{max} = \beta$ . The  $45^\circ$  inclined lines present the equation (3.23), possible (but maybe not practically significant) combinations of the brake forces resulting in the same deceleration value  $\beta$ .



**Figure 25.** Ideal brake force balance for  $\bar{l}_F = 0,52$ ,  $\bar{h} = 0,25$ .

It is obvious, that for a fixed force balance  $-\bar{F}_{xR}/(-\bar{F}_{xF}) = k = \text{constant}$  the ideal utilization of the friction conditions can be used for one specific value of  $\mu_{max}$  only.

If only the wheels of one axle are at the friction limit the corresponding relations for the other longitudinal tyre force due to (3.23), (3.24) are stated with:

friction limit for front locking  $(-\bar{F}_{xF}) = \mu_{max}\bar{F}_{zF}$ :

$$(-\bar{F}_{xR}) = (-\bar{F}_{xF}) \frac{1 - \mu_{max}\bar{h}}{\mu_{max}\bar{h}} - \frac{\bar{l}_R}{\bar{h}}; \quad (3.28)$$

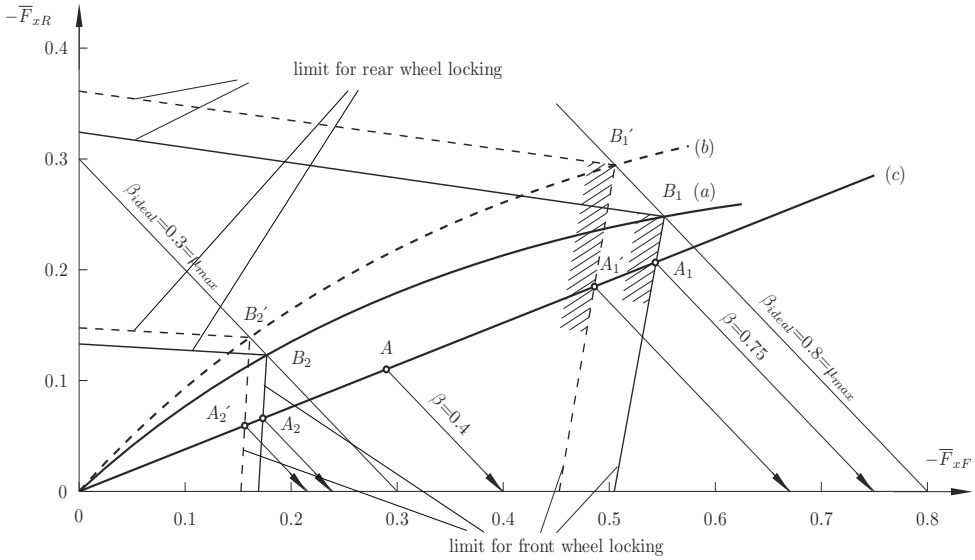
friction limit for rear wheel locking  $(-\bar{F}_{xR}) = \mu_{max}\bar{F}_{zR}$ :

$$(-\bar{F}_{xF}) = -(-\bar{F}_{xR}) \frac{1 + \mu_{max}\bar{h}}{\mu_{max}\bar{h}} + \frac{\bar{l}_F}{\bar{h}}. \quad (3.29)$$

So these straight lines (3.28),(3.29) in a  $(-\bar{F}_{xF}), (-\bar{F}_{xR})$  diagram characterize the limits for a proper utilization of  $\mu_{max}$  at front/rear wheels, see e.g. [11].

Now the information by (3.27), (3.28), (3.29) can be summarized in one brake performance diagram, see e.g. [14], Fig. 26. For the understanding of this figure first have a look at the curves a) and b), which correspond to the ideal brake force curve of Fig.

25, here for two different loading conditions of the same passenger car. Analogously the  $45^\circ$  inclined straight lines correspond to a deceleration with  $\beta = \text{constant}$ . Additionally the limits due to (3.28), (3.29) for front and rear wheel locking for  $\mu_{max} = 0,3$  and  $\mu_{max} = 0,8$  are shown. Furthermore curve c) presents the constructive available constant brake force balance between front and rear wheel  $\bar{F}_{xF}/\bar{F}_{xR} = k$ . With an increase of the brake pressure the actual brake forces are presented by points on this line with increasing distance to the origin. So e.g. point A provides a deceleration value  $\beta = 0,4$  (assuming that  $\mu_{max}$  is large enough).



**Figure 26.** Brake force distribution diagram for a passenger car with 2 loading conditions: a)  $\bar{l}_F = 0,47, \bar{h} = 0,2$ ; b)  $\bar{l}_F = 0,52, \bar{h} = 0,19$ ; design brake balance c)  $\bar{F}_{xF}/\bar{F}_{xR} = k = 2,63$ .

Point  $B_1$  in Fig. 26, the intersection of the locking of front and rear wheels lines, represents one point of the ideal brake force balance with all wheels using the  $\mu_{max} = 0,8$  at the same time (which would be an equivalent of  $\beta_{ideal} = 0,8$ ). As it can be noticed by point  $A_1$  the actual installed brake balance (line c)) reaches the limit line for front wheel locking at  $A_1$ . This means, that with increasing brake pressure the front wheels will first use the available  $\mu_{max}$ , while the rear wheels do not. Consequently the achieved deceleration value  $\beta \cong 0,75$  is smaller than the ideal possible value of  $\beta_{ideal}$ . For the loaded vehicle, curve b), the corresponding points are  $A_1', B_1'$ . Especially in this case the difference between  $\beta$  and  $\beta_{ideal}$  is significantly larger than for curve a). Also included in Fig. 26 are the same results for the low value  $\mu_{max} = 0,3$  with the indices 2.

Preferably the brake balance curve c) is chosen in such a way, that in case of increasing brake pressure the front wheels will reach  $\mu_{max}$  and lock first thereby keeping the vehicle directionally stable (no introduction of a spin motion).

The shaded areas near  $A_1, B_1$  and  $A_1', B_1'$  want to indicate, that a good ABS system



may change the brake forces near the borders of front (and rear) wheel locking. Thereby no fixed brake force balance is given.

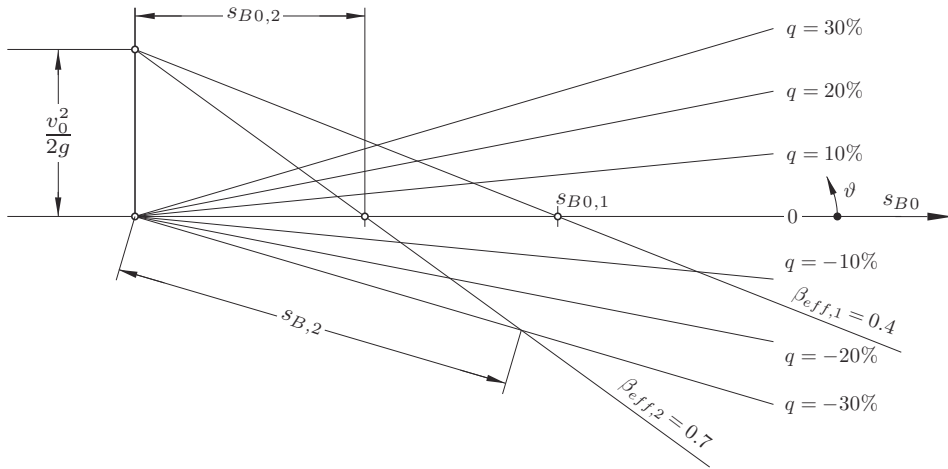
To get an impression of the braking distances also for inclined roads an effective deceleration coefficient can be defined by (3.1), (3.2),

$$\beta_{eff} = \frac{-(2F_{xF} + 2F_{xR})}{2F_{zF} + 2F_{zR}} = \frac{-ma_x + W_L + mg \sin \vartheta}{mg \cos \vartheta - W_Z} \tag{3.30}$$

with  $a_x < 0$ ,  $F_{xF}, F_{xR} < 0$ . For horizontal road  $\vartheta = 0$  and neglecting the aerodynamics  $W_L \cong W_Z \cong 0$  it holds  $\beta_{eff} = \beta$ . Assuming further on also no aerodynamic influences, the stopping distance from an initial velocity of  $v_0$  is calculated by

$$s_B = \frac{v_0^2}{2g(\beta_{eff} \cos \vartheta + \sin \vartheta)}, \quad \text{for } \vartheta = 0 : \quad s_{B0} = \frac{v_0^2}{2g\beta_{eff}}. \tag{3.31}$$

For different grades the stopping distances can be extracted from Fig. 27. The intersections of the  $\beta_{eff,i}$  lines with the horizontal axis provides the stopping distance for  $\vartheta = q = 0$  while the other intersections e.g. with  $\tan \vartheta = q = -30\%$  provides the stopping distance  $s_{B,2}$  for a high friction road surface comparable to  $\mu_{max} \sim 0,7$ .  $s_{B,2}$  is nearly twice the distance of  $s_{B0,2}$  for  $\vartheta = 0$ .



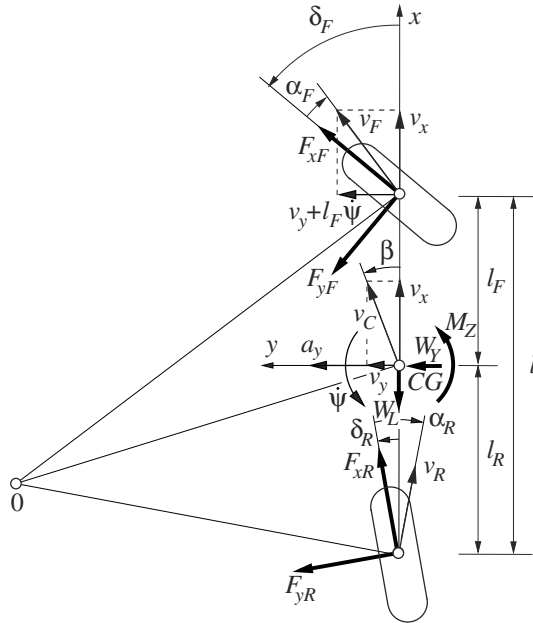
**Figure 27.** Stopping distances  $s_{Bi}$  for different grades  $q$  and two different friction conditions, initial velocity  $v_0$ .

### 4 Lateral Dynamics, Vehicle Handling

The main handling characteristics of a (passenger) car can be derived by a linearized 2-wheel model, [1], [11], [12]. This model, see Fig. 28, is based on the following assumptions.

- The model is built by one substitutive wheel per axle (representing tyre and suspension characteristics) as a plane model correspondingly with the CG in the plane of motion

- All considered angles in Fig. 28 are small and therefore trigonometric functions are linearized
- The tyre forces of the substitutive wheels are linear functions of the side slip angles  $\alpha_i$  only
- No influences of the longitudinal forces will be considered for the vehicle handling.



**Figure 28.** Scheme of plane 2-wheel model.

This model has  $2DOF$ : besides the yaw rate  $r = \frac{d\psi}{dt}$  the side slip angle  $\beta$  of the vehicle or the lateral velocity  $v_y$  can be used. In Fig. 28 there are (substitutive) steering angles  $\delta_F, \delta_R$  at the front and rear wheel; so a vehicle with additional rear wheel steering can be investigated too.  $W_L, W_Y, M_Z$  are the aerodynamic forces and the moment according to chapter 3.2. The notations of the tyre forces  $F_{xi}, F_{yi}$  correspond to chapter 2. The introduced kinematics are self explaining.

Having in mind the simplifications of this 2-wheel model, the possible areas for analyzing the handling properties include

- steady state cornering and steering behaviour
- frequency response
- stability for straight ahead driving
- influence of ambient wind
- basic configuration for control design and for driver model concepts

In this part the last two application areas will not be considered.

#### 4.1 System equations

By Fig. 28 the equations of motion can be derived, see e.g. [16]. For small angles  $\delta_i$  and an assumed forward velocity  $v_x \cong v_c = \text{constant}$  it provides for the plane motion 3 equations:

$$\begin{aligned} 0 &= (F_{xF} - F_{yF}\delta_F) + (F_{xR} - F_{yF}\delta_R) - W_L \\ ma_y &= (F_{xF}\delta_F + F_{yF}) + (F_{xR}\delta_R + F_{yR}) + W_Y \\ I_Z\ddot{\psi} &= (F_{xF}\delta_F + F_{yF})l_F - (F_{xR}\delta_R + F_{yR})l_R + M_Z. \end{aligned} \quad (4.1)$$

To emphasis the main influences and provide simpler analytical expression often the aerodynamics and, as mentioned before, the longitudinal forces are neglected rendering for the lateral dynamics the 2 equations:

$$ma_y = F_{yF} + F_{yR} \quad (4.2)$$

$$I_Z\ddot{\psi} = F_{yF}l_F - F_{yR}l_R \quad (4.3)$$

where the forces of the substitutive tyres are described (analogue to equation (2.3)) by

$$F_{yi} = C_i\alpha_i, \quad i = R, F. \quad (4.4)$$

The corresponding kinematics, Fig. 28, using the body fix reference frame, provides for the lateral acceleration

$$a_y = \dot{v}_y + \dot{\psi}v_x \quad (4.5)$$

or

$$a_y = v_x(\dot{\beta} + \dot{\psi})$$

when using the vehicle side slip angle (for assumed constant  $v_x$ )

$$\beta = \frac{v_y}{v_x}, \quad \dot{\beta} = \frac{\dot{v}_y}{v_x}.$$

For the side slip angles of the wheels the relations can be written with

$$\begin{aligned} \alpha_F &= \delta_F - \frac{v_y + l_F\dot{\psi}}{v_x}, \\ \alpha_R &= \delta_R - \frac{v_y + l_R\dot{\psi}}{v_x} \end{aligned}$$

or

$$\begin{aligned} \alpha_F &= \delta_F - \beta - \frac{l_F\dot{\psi}}{v_x}, \\ \alpha_R &= \delta_R - \beta + \frac{l_R\dot{\psi}}{v_x}. \end{aligned} \quad (4.6)$$

Now using the equations (4.2) till (4.6) the dynamic behaviour can be presented with a system of linear differential equations of first order with constant coefficients (with  $r = \dot{\psi}$ ):

$$\dot{\underline{x}} = \mathbf{F}\underline{x} + \mathbf{G}\delta \quad (4.7)$$

with

$$\underline{x} \begin{bmatrix} v_y \\ \psi \end{bmatrix} = \begin{bmatrix} v_y \\ r \end{bmatrix}, \quad \underline{\delta} = \begin{bmatrix} \delta_F \\ \delta_R \end{bmatrix},$$

$$\mathbf{F} = \begin{bmatrix} -\frac{C_F + C_R}{mv_x} & -\frac{(l_F C_F - l_R C_R)}{mv_x} + v_x \\ -\frac{(l_F C_F - l_R C_R)}{I_Z v_x} & -\frac{l_F^2 C_F + l_R^2 C_R}{I_Z v_x} \end{bmatrix},$$

$$\mathbf{G} = \begin{bmatrix} \frac{C_F}{m} & \frac{C_R}{m} \\ \frac{l_F C_F}{I_Z} & \frac{l_R C_R}{I_Z} \end{bmatrix}.$$

These equations can also be written with second order differential equations, that for same cases offer better direct insight into problems. For a vehicle without rear wheel steering ( $\delta_R = 0$ ), which will be considered further on, the equations are

$$\begin{aligned} \ddot{\beta} + 2K_1 \dot{\beta} + K_2 \beta &= \frac{C_F}{mv_x} \dot{\delta}_F - \frac{C_F(l_F m v_x^2 - C_R l_R l)}{I_Z m v_x^2} \delta_F, \\ \ddot{r} + 2K_1 \dot{r} + K_2 r &= \frac{l_F C_F}{I_Z} \dot{\delta}_F + \frac{C_F C_R l}{I_Z m v_x} \delta_F \end{aligned} \quad (4.8)$$

with

$$\begin{aligned} K_1 &= \frac{I_Z(C_R + C_F) + m(C_F l_F^2 + C_R l_R^2)}{2I_Z m v_x}, \\ K_2 &= \frac{l^2 C_F C_R + (C_R l_R - C_F l_F) m v_x^2}{I_Z m v_x^2}. \end{aligned}$$

## 4.2 Steady State Cornering

Quite a number of characteristic values for the vehicle behaviour is derived from the steady state cornering of the car and also defined with the SAE J670e terminology, e.g. [1].

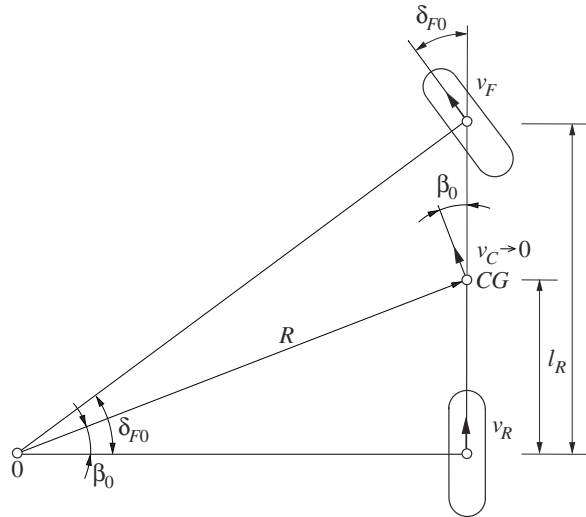
The steady state trim conditions, with cornering radius  $R$  and  $v_x \cong v_c$  are given by

$$\beta_{st} = r_{st} = \text{constant}, \quad a_{y,st} = \frac{v_x^2}{R}. \quad (4.9)$$

As limit values for a condition  $v_x \cong v_c \rightarrow 0$ , where no lateral forces and thereby no wheel side slip is necessary, the steering angle  $\delta_{F0}$  and the vehicle side slip angle can be directly determined by Fig. 29:

$$\beta_o = \frac{l_R}{R}, \quad \delta_a = \delta_{F0} = \frac{l}{R} \quad \text{or} \quad \delta_a = \frac{l}{v_x^2} a_{y,st}. \quad (4.10)$$

In (4.10) there is also introduced the Ackermann angle  $\delta_a$  (also defined not only for  $v_x \rightarrow 0$ ).



**Figure 29.** Geometric conditions for steady state cornering with  $v_c \rightarrow 0$ .

Using the equations of motion, e.g. (4.8), the steady state values of the necessary steering angle  $\delta_{F,st}$  and the corresponding side slip angle become

$$\delta_{F,st} = \delta_{F0} + \frac{C_R l_R - C_F l_F}{C_R C_F l} m a_{y,st}, \quad (4.11)$$

$$\beta_{st} = \beta_0 - \frac{l_F}{C_R l} m a_{y,st} \quad (4.12)$$

and very clearly show the changes with respect to the condition  $v_c \rightarrow 0$ .

The so called "steering behaviour" is defined by the understeer/oversteer gradient (Eigenlenkgradient) EG based on the change of  $(\delta_{F,st} - \delta_a)$

$$EG = \frac{1}{i_S} \frac{\partial \delta_{H,st}}{\partial a_{y,st}} - \frac{\partial \delta_a}{\partial a_{y,st}} \quad (4.13)$$

where the steady state steering wheel angle  $\delta_{H,st}$  is used. With the steering gear ratio  $i_S$  it holds

$$\delta_{H,st} = i_S \cdot \delta_{F,st}. \quad (4.14)$$

Then the steering angle (4.11) can be expressed by

$$\frac{\delta_{H,st}}{i_S} = \delta_{F,st} = \delta_a + EG a_{y,st} \quad (4.15)$$

with

$$EG = \frac{m(C_R l_R - C_F l_F)}{l C_R C_F}. \quad (4.16)$$

Obviously by (4.15) the gradient EG determines (since  $a_{y,st} > 0$ ) how the steering angle has to be changed compared to the Ackermann angle  $\delta_a$ . Consequently the steering properties are characterized with

$$\begin{aligned} EG &> 0 && \text{understeer} \\ EG &= 0 && \text{neutral steer} \\ EG &< 0 && \text{oversteer.} \end{aligned} \tag{4.17}$$

With the kinematics shown in Fig. 30

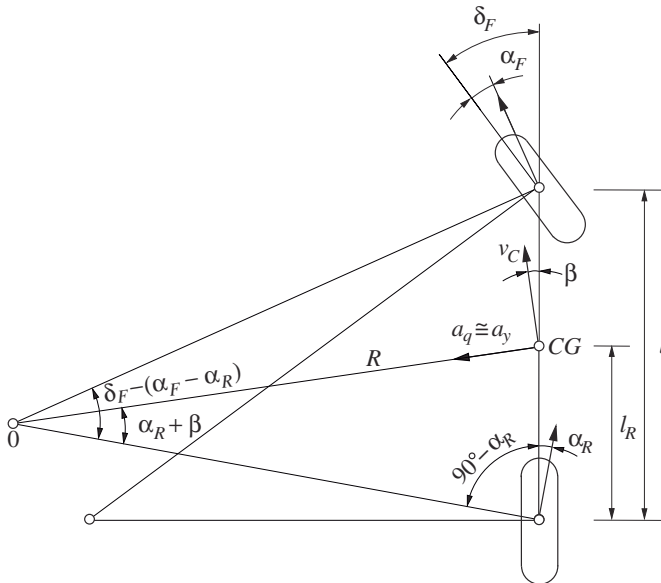
$$\delta_{F,st} = \frac{l}{R} + (\alpha_F - \alpha_R) = \delta_a + (\alpha_F - \alpha_R)$$

and/or with the dynamic (4.2), (4.3), (4.4) for steady state cornering (small angles):

$$\alpha_F = \frac{l_R}{lC_F} ma_y, \quad \alpha_R = \frac{l_F}{lC_R} ma_y$$

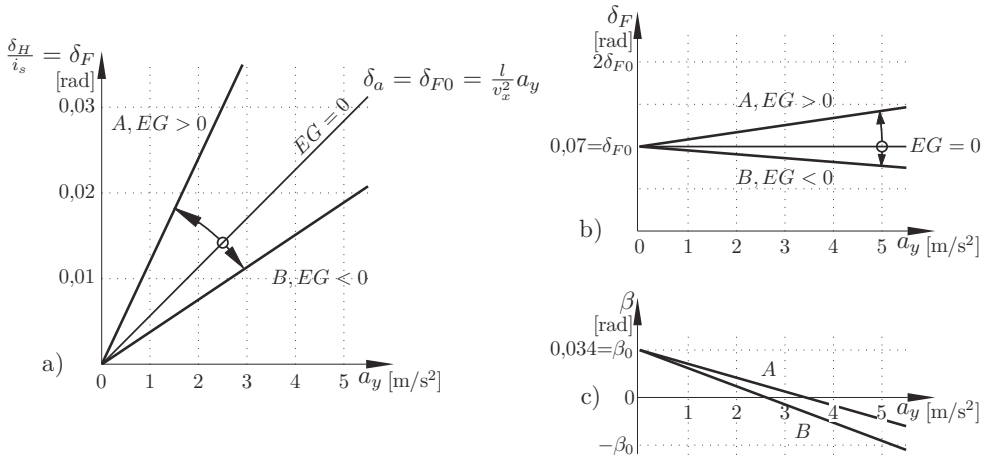
a direct connection can be established between EG and the difference of the side slip angles

$$(\alpha_F - \alpha_R) = EG a_{y,st}. \tag{4.18}$$



**Figure 30.** Kinematics for steady state cornering.

To characterize the steering properties (and the corresponding side slip angle) as functions of the lateral acceleration two kinds of diagrams are used, see e.g. [14], Fig. 31. The necessary vehicle data are given in Table 1.



**Figure 31.** Steady state steering characteristics, data corresponding to Table 1. a) for  $v = \text{constant} = 80 \text{ km/h}$  b) for  $R = \text{constant} = 40 \text{ km/h}$  c) side slip angles of the cars corresponding to b).

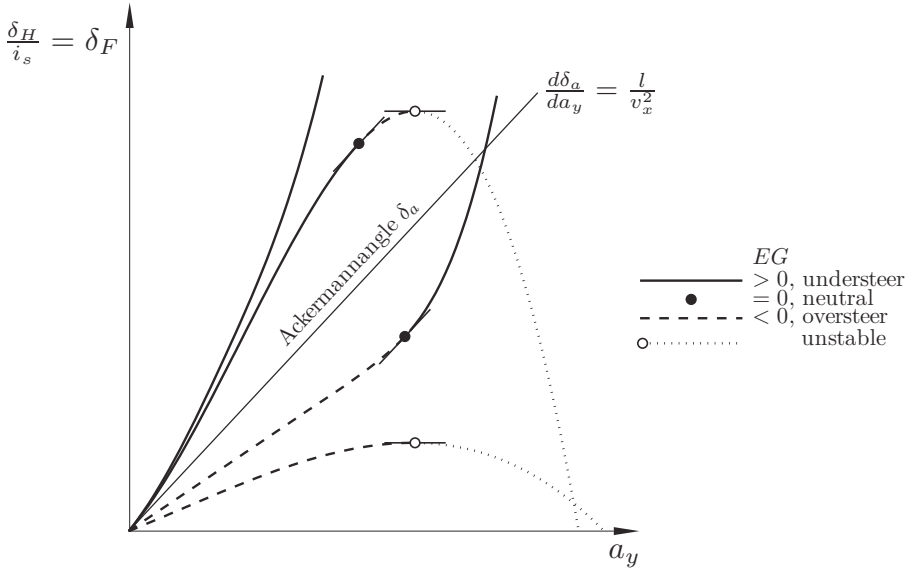
Vehicle	A	B
$m$	1900 kg	
$I_Z$	2900 $\text{kgm}^2$	
$l_F$	1,44 m	
$l_R$	1,36 m	
$C_F$	90 000 N	60 000 N
$C_R$	80 000 N	110 000 N
$EG$	$-1,95 \cdot 10^{-3} \text{ s}^2 \text{ m}^{-1}$	$+6,51 \cdot 10^{-3} \text{ s}^2 \text{ m}^{-1}$
Steering charact.	oversteer	understeer

**Table 1.** Vehicle data for two passenger cars with different steering characteristics.

In a more general case of a nonlinear behaviour the steering behaviour characteristic may change with the lateral acceleration like shown in Fig. 32 with the SAE J670e terminology, [1].

Special values of velocities, that also help to characterize the handling behaviour, are defined using the steady state yaw velocity gain, [14]

$$G_{r,st} = \frac{r}{\delta_H/i_s} \Big|_{st} = \frac{v_x}{l + EG \cdot v_x^2}. \quad (4.19)$$



**Figure 32.** General case steering behaviour definition by the EG gradient according to the SAE J670e terminology.

For an understeer vehicle a characteristic velocity is defined by

$$\begin{aligned} \frac{\partial G_{r,st}}{\partial v_x} \Big|_{v_x=v_{ch}} &= \frac{l - EGv_{ch}^2}{(l + EGv_{ch}^2)^2} = 0; \\ v_{ch}^2 &= \frac{l}{EG}, \quad EG > 0 \end{aligned} \tag{4.20}$$

while for an oversteer vehicle a critical velocity can be given with

$$\begin{aligned} G_{r,st} \Big|_{v_x=v_{crit}} &\rightarrow \infty, \\ v_{crit}^2 &= -\frac{l}{EG}, \quad EG < 0. \end{aligned} \tag{4.21}$$

By this yaw velocity gain the sensitivity of the vehicle to a steering input can be interpreted. Especially for high velocities it should not be too large. Fig. 33 shows, that the understeer vehicle will have a good moderate response while especially the oversteer vehicle has quite strong responses at higher velocities near the critical one.

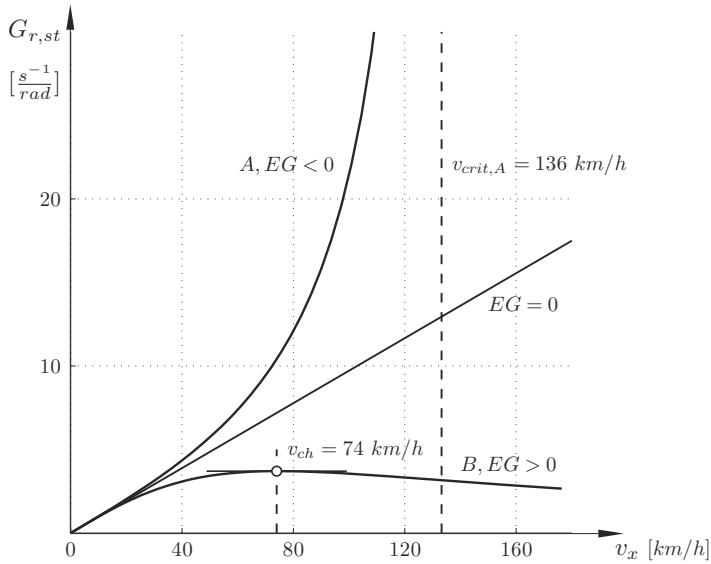
Analogously a steady state acceleration gain can be considered. According to the relation

$$a_{y,st} = r_{st} \cdot v_x$$

the acceleration response to the steering input becomes

$$\frac{a_y}{\delta_H/i_s} \Big|_{st} = G_{ay,st} = G_{r,st} \cdot v_x = \frac{v_x^2}{l + EG \cdot v_x^2}. \tag{4.22}$$





**Figure 33.** Behaviour of oversteer ( $EG < 0$ ), neutral ( $EG = 0$ ) and understeer ( $EG > 0$ ) vehicle with respect to the static yaw velocity gain; vehicle data for  $A, B$  according table 1.

Corresponding to Fig. 33 the Fig. 34 for the steady state acceleration gain as function of the velocity shows that only the understeer vehicle will have a moderate response for higher velocities.

It needs to be mentioned that the definitions of oversteer and understeer are derived and based on the steady state behaviour of the car. An application of this terminology for non steady state conditions should be avoided or at least handled with special care.

### 4.3 Non Steady State Behaviour

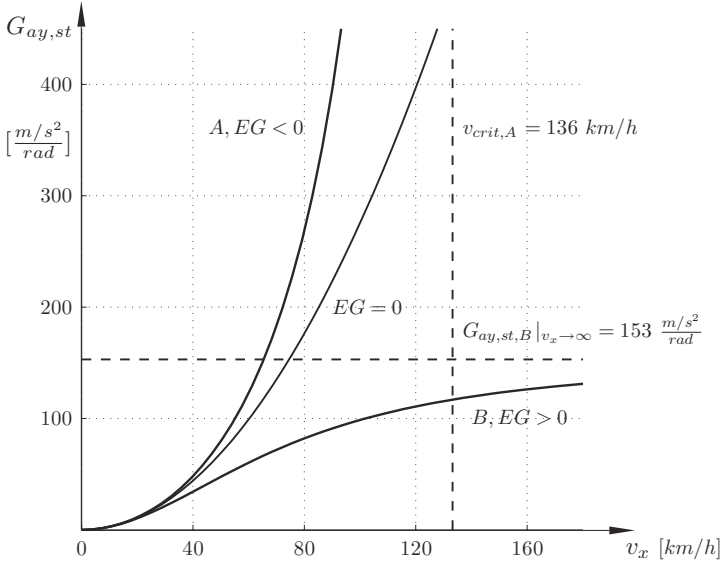
As before the following basic characterization of special non steady state driving conditions are derived by the 2-wheel model and definitions already used for the steady state cornering chapter 4.2.

The stability behaviour for small perturbations (straight ahead driving,  $\delta_F \equiv 0$ ) is determined by the left hand side of (4.8), the homogenous differential equations that can be written in the form

$$\begin{aligned} \ddot{x}_i + 2D\omega_0\dot{x}_i + \omega_0^2x_i &= 0 & x_i &= \beta, r \\ \omega_0^2 &= K_2, & D\omega_0 &= K_1. \end{aligned} \quad (4.23)$$

Since always  $2D\omega_0 = K_1 > 0$  as can be noticed by (4.8), the stability of the system is determined by  $\omega_0^2$ . With using  $EG$  it can be written

$$\omega_0^2 = K_2 = \frac{C_F C_R l^2}{I_Z m v_x^2} \left( 1 + \frac{EG v_x^2}{l} \right). \quad (4.24)$$



**Figure 34.** Static lateral acceleration gain corresponding to Fig. 33

For an understeer vehicle ( $EG > 0$ ) this can be expressed by  $v_{ch}$ , see (4.20)

$$\omega_{o|u}^2 = \frac{C_F C_R l^2}{I_Z m v_x^2} \left( 1 + \frac{v_x^2}{v_{ch}^2} \right); \quad v_{ch}^2 = \frac{l}{EG} > 0 \quad (4.25)$$

and clearly shows that always  $\omega_{o|u}^2 > 0$  is valid and therefore the vehicle is stable for all velocities  $v_x$ .

On the contrary for the oversteer vehicle ( $EG < 0$ ) equation (4.24) can be written using the critical velocity, see (4.21) with

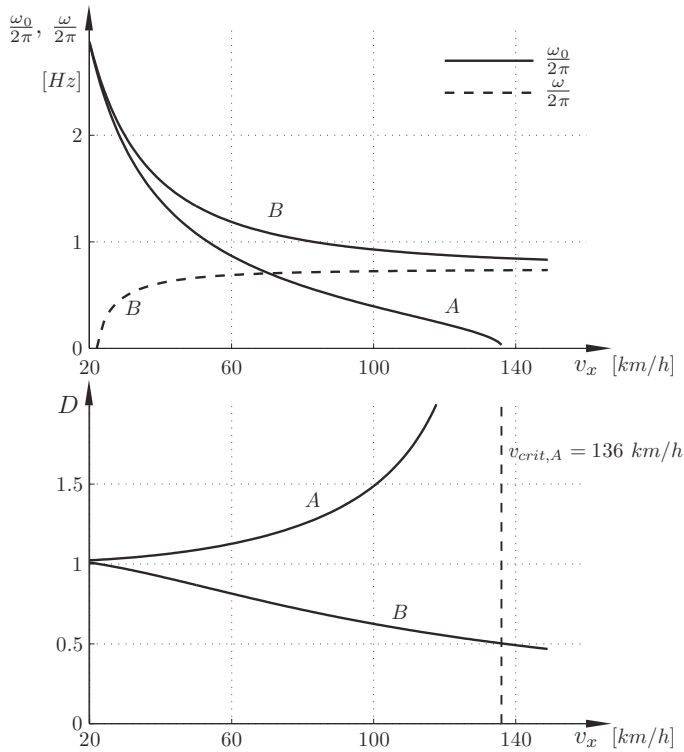
$$\omega_{o|o}^2 = \frac{C_F C_R l^2}{I_Z m v_x^2} \left( 1 - \frac{v_x^2}{v_{crit}^2} \right); \quad v_{crit}^2 = -\frac{l}{EG} > 0. \quad (4.26)$$

For  $v_x > v_{crit}$  then  $\omega_{o|o}^2 < 0$  and the system is unstable. The critical velocity  $v_{crit}$  is the limit for a possible regular driving with respect to perturbations – compare also Fig. 33, 34. The eigenfrequencies and the damping coefficient  $D$  for the over- and understeer vehicle, Table 1, are shown in Fig. 35. Thereby it is differentiated between the eigenfrequency of the undamped system  $\omega_0$ , and  $\omega$  of the damped system

$$\omega^2 = \omega_0^2 (1 - D^2).$$

In Fig. 35 the consequence of  $v_{crit}$  for vehicle A becomes obvious once again. For vehicle A no  $\omega$  can be given, because it has an aperiodic behaviour.

A further characterization of the handling behaviour is the response of the vehicle to



**Figure 35.** Eigenfrequencies and damping coefficient  $D$  for an oversteer and understeer vehicle according to Table 1.

harmonic steering inputs. Especially for frequencies often used with normal driving (up to about 1 Hz) the influences of frequency and to some extent also driving velocities should be very limited for convenient handling.

The yaw velocity frequency response to a harmonic steering excitation with frequency  $\nu$  is characterized by the complex transfer function, see e.g. [14]

$$G_r(i\nu) = \left( \frac{r}{\delta_F} \right) \Big|_{i\nu} = G_{r,st} \frac{1 + T_Z \cdot (i\nu)}{1 + \frac{2D}{\omega_o} \cdot (i\nu) - \frac{\nu^2}{\omega_o^2}} \quad (4.27)$$

with

$$T_Z = \frac{mv_x l_F}{C_R l}.$$

Thereby the steady state yaw gain (4.19) is used.

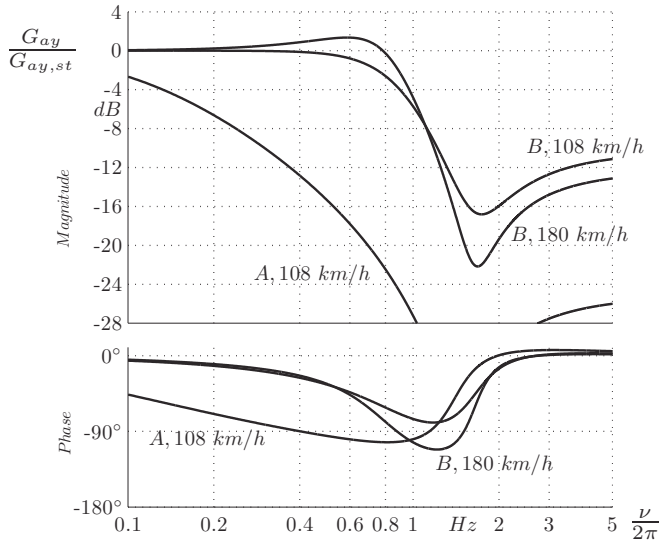
For the lateral acceleration the frequency response is expressed by using  $a_y = v_x(r+\beta)$  and the steady state response (4.22):

$$G_{ay}(i\nu) = \left( \frac{a_y}{\delta_F} \right) \Big|_{i\nu} = G_{ay,st} \frac{1 + T_1 \cdot (i\nu) - T_2 \nu^2}{1 + \frac{2D}{\omega_o} \cdot (i\nu) - \frac{\nu^2}{\omega_o^2}} \quad (4.28)$$

with

$$T_1 = \frac{l_R}{v_x}, \quad T_2 = \frac{I_z}{C_{Rl}}.$$

Once again the comparison of the oversteer vehicle *A* and the understeer vehicle *B*, Table 1, in Fig. 36 shows the distinctive different behaviour for the normalized lateral acceleration gain. For the oversteer vehicle *A* there is no result for 180 km/h because of a  $v_{crit} = 136\text{km/h}$ . Remarkably for vehicle *A* is not only the relative strong dependency of the gain as function of excitation frequency  $\nu$  but also the delay of the response due to the phase angle compared to the understeer vehicle *B*, [17].

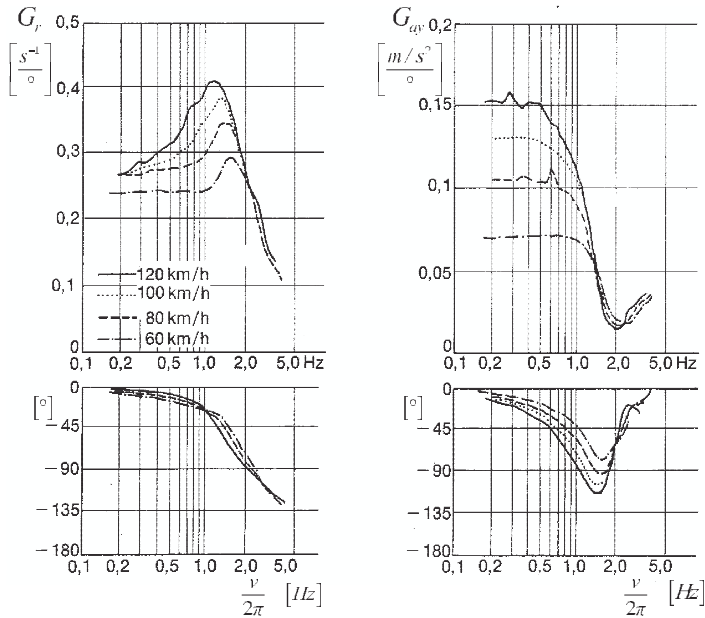


**Figure 36.** Normalized acceleration frequency response of oversteer vehicle *A* and understeer vehicle *B*, data table 1.

Measured frequency responses of an understeer vehicle ( $EG = 0,062$ ;  $v_{crit} = 76$  km/h) very similar to vehicle *A*, Table 1, are shown in Fig. 37, [14]. The relative small phase angles till about 1 Hz indicate a quick response of the vehicle to the steering input like expected by the driver. Moreover there is no much reduced response in the lower frequency range.

An other test for the transient vehicle behaviour that to some extent can be analyzed by the 2-wheel model is the steering step input. The car moves with constant velocity straight ahead and the the driver initiates a steering input at the steering wheel as fast as possible (values larger than  $500^\circ/s$  are noticed) with the target to achieve a special yaw velocity or steady state lateral acceleration. Fig. 38 shows the corresponding idealized steering characteristics, normalized with respect to the desired limit value. The vehicle response, here also shown in normalized form, is analyzed using the given standardized time intervals and the overshoot, [14], [10].

Measured examples and recommended limits defined by ESV (Experimental Safety Vehicle), [18], are presented in Fig. 39. The steering step input was performed with



**Figure 37.** Measurements of yaw velocity and lateral acceleration gain of a vehicle similar to vehicle *A*, Table 1.

$d\delta_H/dt \cong 500^0/s$  and the final steady state cornering has a lateral acceleration of  $a_{y,st} \cong 0,4g$ .

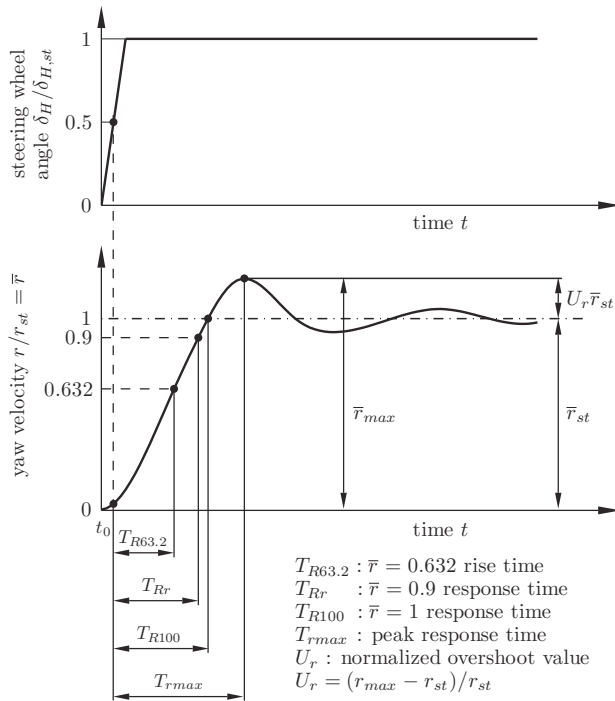
## 5 Advanced Vehicle Models

The area of this topic is nearly unlimited because of the more detailed modeling of the vehicle itself, additional modeling features need to be added or even integrated due to e.g. control aspects, driver behaviour, stress loading of suspension components etc. Consequently only some of the interesting possibilities can be presented.

Once again considering some general aspects of modeling, Fig. 40 wants to show that today much of the circle from problem formulation to simulation results is covered by available software. The main and essential steps for the simulation of an advanced system are the physical modeling and the interpretation of the simulation results and their validation with respect to the problem formulation. Experience and team work are a demand for success, especial the problem oriented modeling including the knowledge of the necessary model parameters!

For an increase in model complexity and expansion the following listing wants to provide a possible overview:

- Linearized models with linear tyre behaviour or without slip
- Special handling models with roll axes and nonlinear lateral tyre behaviour
- 4-wheel models with more or less complex tyre characteristics



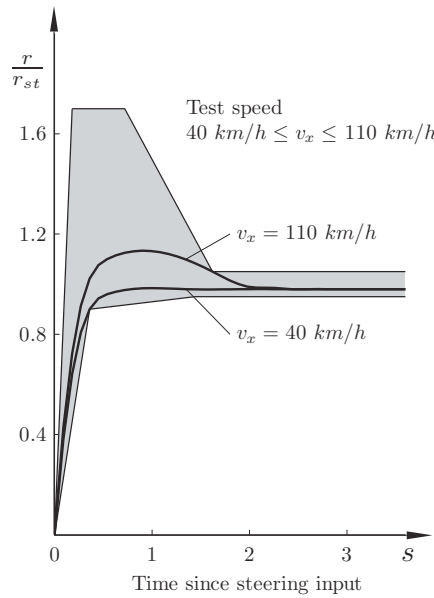
**Figure 38.** Steering step input and vehicle yaw response; principal diagram.

- 3D-Models with separate component descriptions, complex tyre characteristics
- MBS vehicle models with complex tyre description for flat road surface
- MBS vehicle models with structural tyre description for contact with structured road surface.
- Vehicle system model: vehicle model, control systems, driver model

As told before with linearized models the longitudinal, lateral and vertical dynamics can be considered separately. As a completion to the models given by Fig.15 and 28, Fig. 41 shows a plane passenger car model, which is commonly used for ride and comfort investigations, e.g. [12], [20].

In recent times special handling models with roll axes and/or 4-wheel models with nonlinear tyre characteristic are used for a better analysis of passenger car properties based on driving test results. The combination of somewhat simplified but nonlinear vehicle model with measurement information opens the possibility for a better interpretation of e.g. steering characteristics or yaw gain response, [21].

A typical 3D-vehicle model is presented by Fig. 42, [8]. It is an integration of more or less separate components using proper interconnections. So e.g. the relative movement of a wheel with respect to the car body is based on the relative travel of the wheel center and the steering angle  $\delta_{S1}$  at the steering axle of the wheel without details of the connections "wheel" to car body or to the steering system (travel  $M_{O1}$  till  $M_{S1}$ ).



**Figure 39.** Measured yaw responses to steering step inputs with final lateral acceleration  $a_{y,st} = 0, 4g$  and recommended performance region.

The later is modeled separately with  $\delta_{S1}$  as the interface. Similar considerations hold for the drive train – wheel connection. Elastokinematic suspension aspects are added by characteristic functions with the tyre forces as input quantities.

By using the 3D-kinematic description of the wheel center position and the wheel attitude (coordinate system  $x_1, y_1, z_1$  and  $M_{\delta_1}$ ) the local contact between wheel and road surface can be established, Fig. 43. Finally this provides the tyre deflection, the normal tyre force  $F_{z_1}$  and the side slip angle responsible for  $F_{y_1}$  taking into account the mean local surface structure.

The tyre model has to include the mutual influence of longitudinal and lateral tyre forces and it should be applicable for arbitrary slip values. So at least a model complexity like shown by Fig. 7 has to be used.

The equations of motion for such a kind of model can generally be formulated with a minimal number of independent variables  $z$ , as shown in the example of Fig. 44. Thereby Jacobian matrices  $\mathbf{J}$  and transformation matrices  $\mathbf{A}$  are used for the kinematics.  $\underline{F}_i, \underline{M}_i$ , represent the vector sum of the relevant forces and moments for the Newton and Euler equations.

Figures 45, 46 show typical simulation results of a 3D-vehicle model in comparison with measurements (internal company information), for steady state cornering and an emergency braking out of a steady state cornering with a deceleration of approximately  $7,5 \text{ m/s}^2$ . For this emergency braking the simulation very well reproduces the overall behaviour: the vehicle starts to leave the curve to the outside nearly in a tangential way as can be seen by the lateral acceleration  $a_y$ . The simulation naturally fails to describe

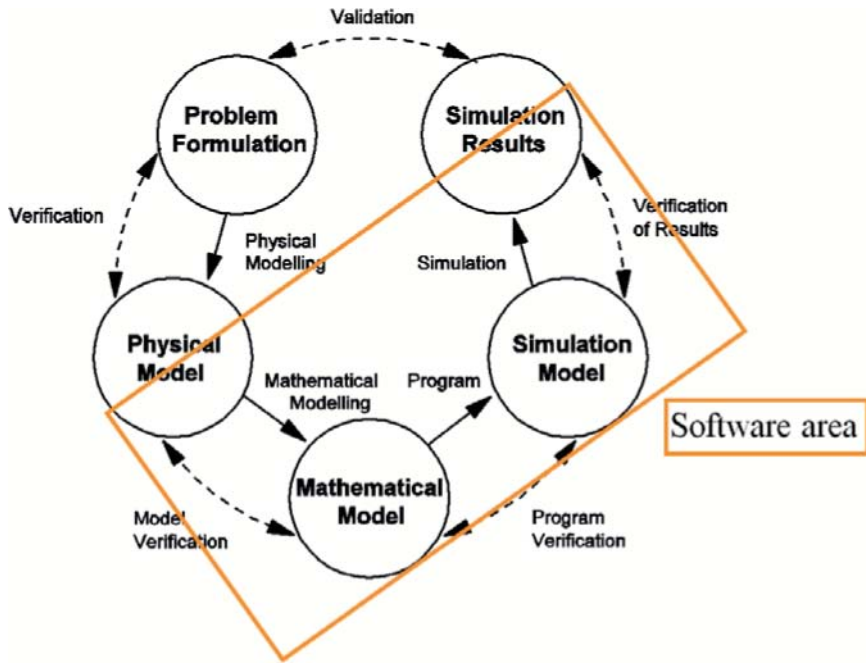


Figure 40. Schema of the modeling process.

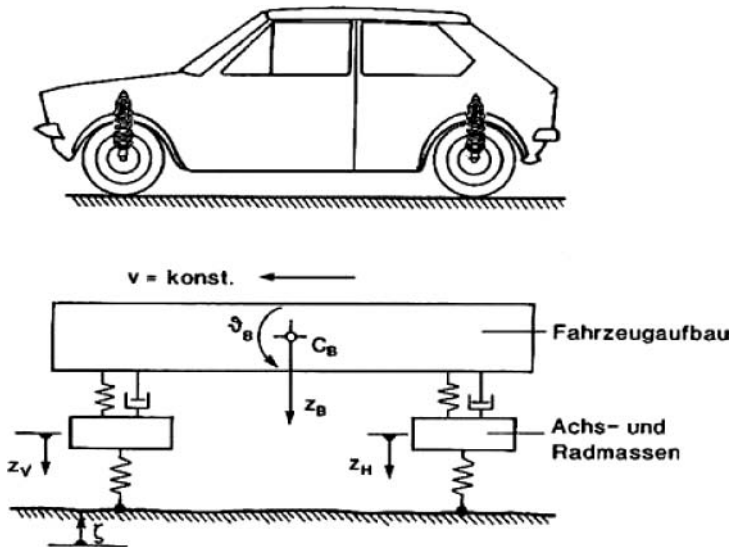


Figure 41. Model for vertical dynamics.



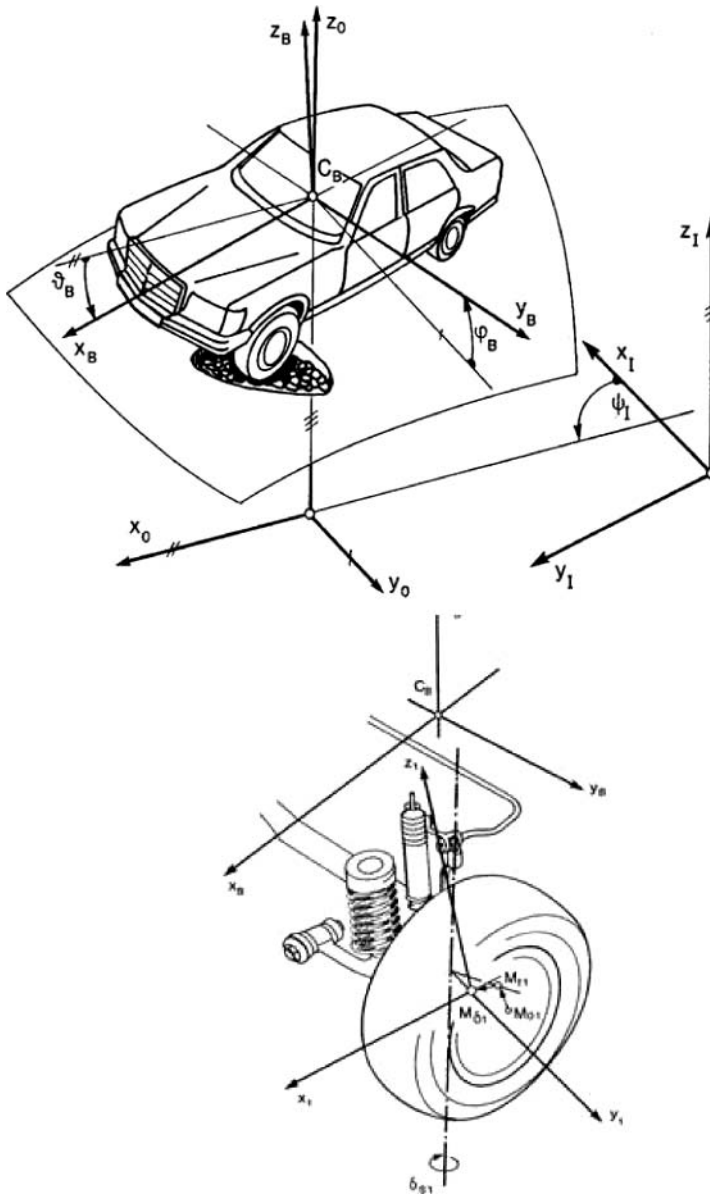
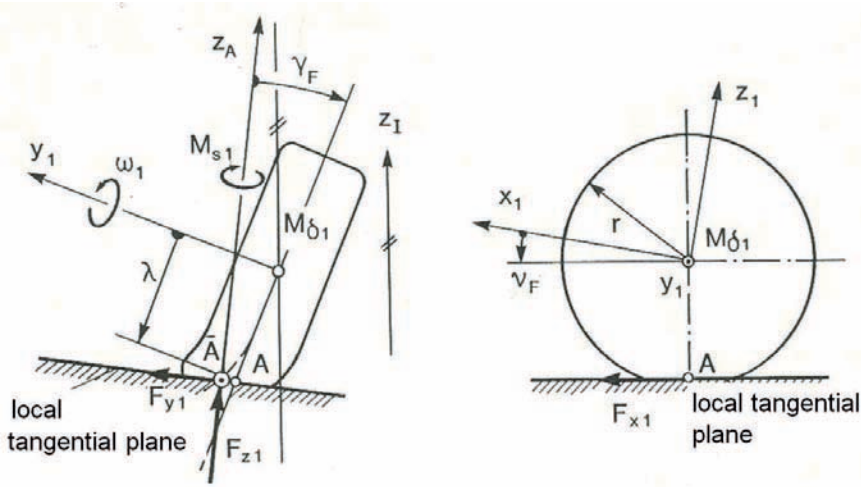


Figure 42. 3D-vehicle and suspension model of the left front wheel.



**Figure 43.** Local tyre road contact taking into account the tyre elasticities, corresponding to Fig. 42.

$$m_i \mathbf{J}_{Ci} \ddot{\mathbf{z}} + m_i \underline{a}_i^* = \underline{F}_i$$

$$\mathbf{I}_i \mathbf{J}_{\omega i} \ddot{\mathbf{z}} + \mathbf{I}_i \underline{\alpha}_i^* + \left( \widetilde{\mathbf{J}}_{\omega i} \dot{\mathbf{z}} + \widetilde{\underline{\omega}}_i^* \right) \mathbf{I}_i \left( \mathbf{J}_{\omega i} \dot{\mathbf{z}} + \underline{\omega}_i^* \right) = \underline{M}_i,$$

$$\begin{aligned} \underline{v}_{i/i} &= \mathbf{A}_{iI} \left( \mathbf{J}_{Ci}(\mathbf{z}, t) \dot{\mathbf{z}} + \underline{v}_i^*(\mathbf{z}, t) \right) \\ &= \mathbf{J}_{Ci/i}(\mathbf{z}, t) \dot{\mathbf{z}} + \underline{v}_{i/i}^*(\mathbf{z}, t), \end{aligned}$$

$$\underline{\omega}_{Ii/i} = \underline{\omega}_{i/i} = \mathbf{J}_{\omega i/i}(\mathbf{z}, t) \dot{\mathbf{z}} + \underline{\omega}_{i/i}^*(\mathbf{z}, t),$$

**Figure 44.** Structure of a typical formulation of the equations of motion for a 3D-model.

details sufficiently. Having a look at the yaw velocity  $\dot{\psi}$  in the region till the beginning of the braking,  $t = 0,5s$ , Fig. 46 indicates that the deviations simulation – measurement are also remarkably influenced by the road condition and the driver input.

For more details in the modeling and a nearer to reality connection of e.g. the wheel suspensions to the car body, the vehicle will be considered as a multibody system (MBS), [23]. The equations of motion will have the structure of a DAE-System like shown in Fig. 47 where the algebraic equations  $g(p) = 0$  result by closed loop constrains when joints are modeled as rigid connection. Today available MBS-software like SIMPACK, ADAMS, etc. do not use single bodies or joints only but already provide whole parameterized subcomponent descriptions which can be integrated to the whole vehicle, see Fig. 48. The picture of a tyre indicates that a tyre model (or even a choice of different models) is available for the investigations.

These subcomponents may include quite a number of structural details, see Fig. 49,

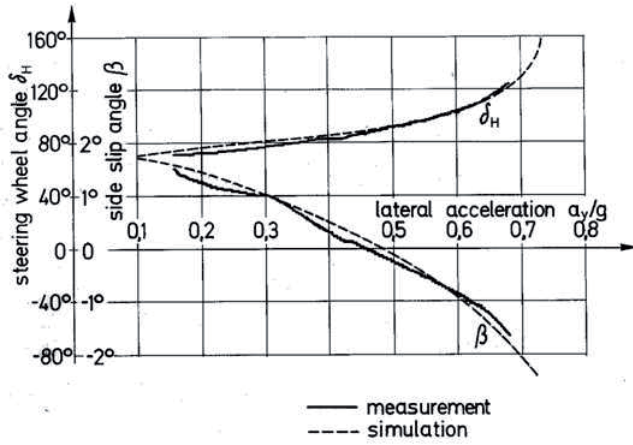


Figure 45. 3D-vehicle model: comparison of simulation results and measurements for steady state cornering,  $R \approx 40\text{m}$ .

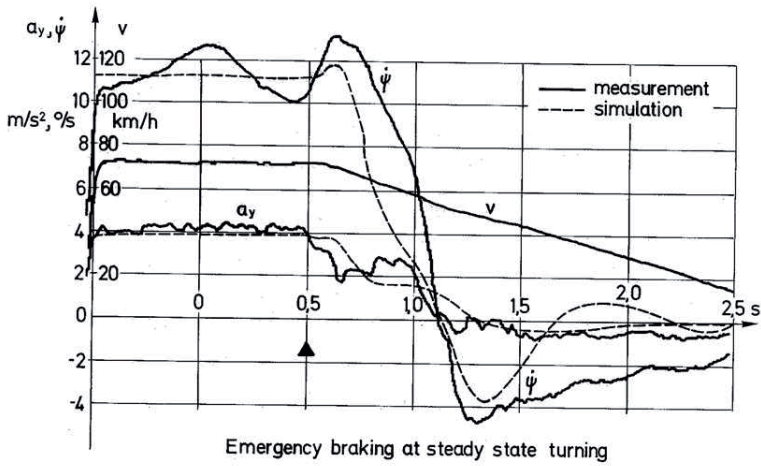
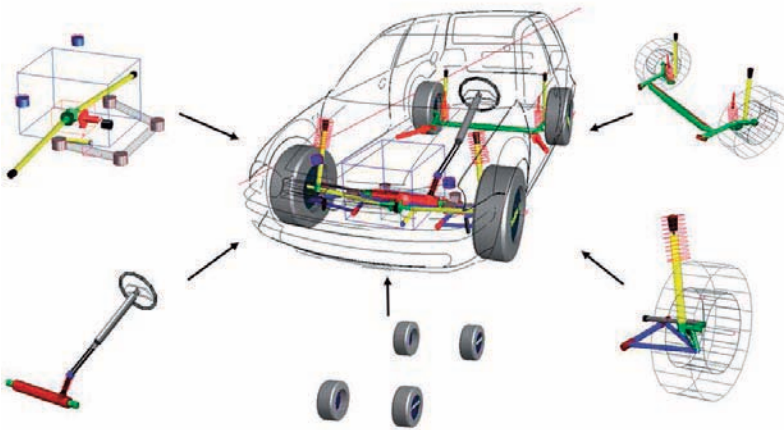


Figure 46. Emergency braking manoeuvre out of steady state cornering.

$$\begin{aligned} \dot{\mathbf{p}}(t) &= \mathbf{T}(\mathbf{p})\mathbf{v}(t) \\ \mathbf{M}(\mathbf{p})\dot{\mathbf{v}} &= \mathbf{f}^*(\mathbf{p}, \mathbf{v}, t) + \mathbf{G}^T(\mathbf{p})\lambda \\ \mathbf{g}(\mathbf{p}) &= \mathbf{0} \\ \dot{\mathbf{g}}(\mathbf{p}, \mathbf{v}) &= \frac{\partial \mathbf{g}(\mathbf{p})}{\partial \mathbf{p}}\mathbf{T}(\mathbf{p})\mathbf{v} = \mathbf{G}(\mathbf{p})\mathbf{v} = \mathbf{0} \\ \ddot{\mathbf{g}}(\mathbf{p}, \mathbf{v}, \dot{\mathbf{v}}) &= \mathbf{G}(\mathbf{p})\dot{\mathbf{v}} + \dot{\mathbf{G}}(\mathbf{p}, \mathbf{v})\mathbf{v} = \mathbf{0} \end{aligned}$$

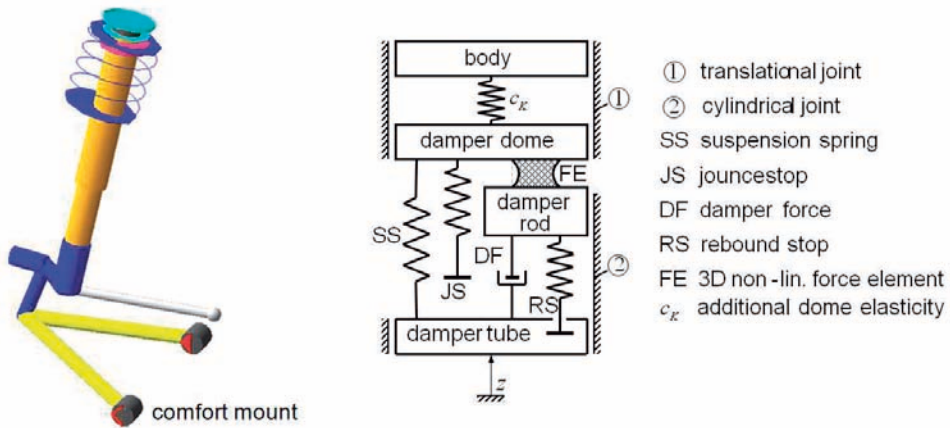
- $\mathbf{p} \in \mathbb{R}^{n_p}$  : generalized coordinates  
 $\mathbf{v} \in \mathbb{R}^{n_p}$  : generalized velocities  
 $\mathbf{M}(\mathbf{p}) \in \mathbb{R}^{n_p \times n_p}$  : generalized mass matrix  
 $\mathbf{f}^*(\mathbf{p}, \mathbf{v}, t) \in \mathbb{R}^{n_p}$  : generalized forces, inertia and gyro effects  
 $\lambda \in \mathbb{R}^{n_z}$  : Lagrange multiplier  
 $\mathbf{g}(\mathbf{p}) \in \mathbb{R}^{n_z}$  : position constraint

**Figure 47.** Principal formulation of equations of motion for a MBS-System.



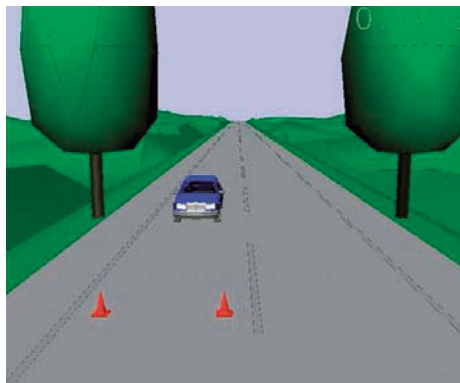
**Figure 48.** Assembling of a passenger car using the MBS-software SIMPACK.

and flexible parts. The description of the flexibilities can be derived by FEM calculations which are then transformed into a kind of shape functions for the MBS simulation.



**Figure 49.** McPherson axle: schematic presentation of the strut with connection to the car body.

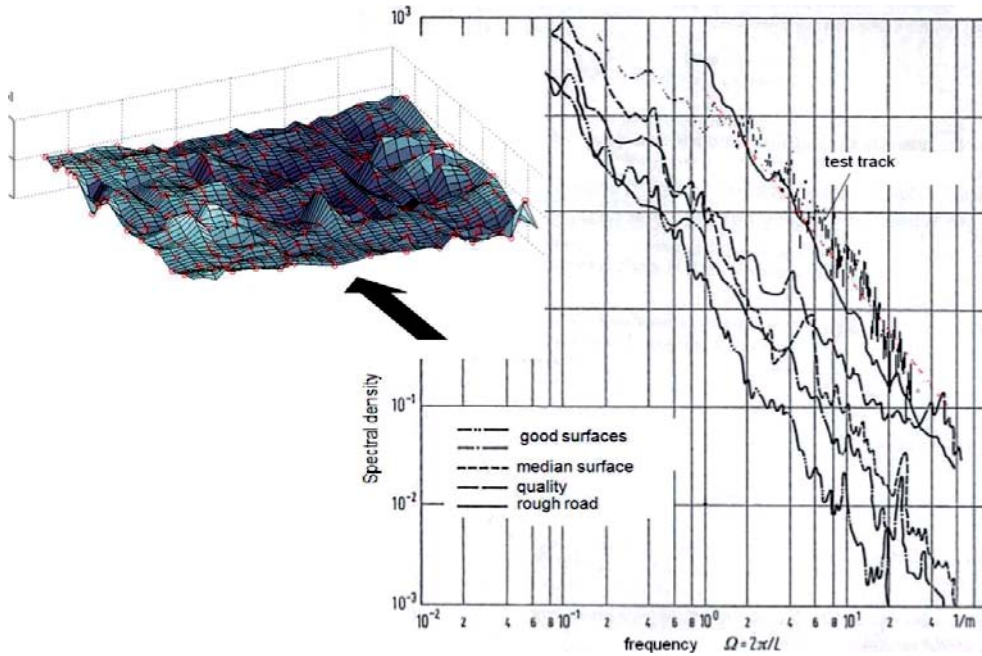
Besides the vehicle itself the MBS-software provides the possibility of a presentation of surrounding features like the road. So as a result of the simulation, animations of a vehicle moving in a stylized world can be presented – sometimes a very essential feature for non-experts to better understand a manoeuvre and the vehicle responses. Fig. 50 provides an impression of an animation surrounding and the passenger car that will perform a lane change manoeuvre.



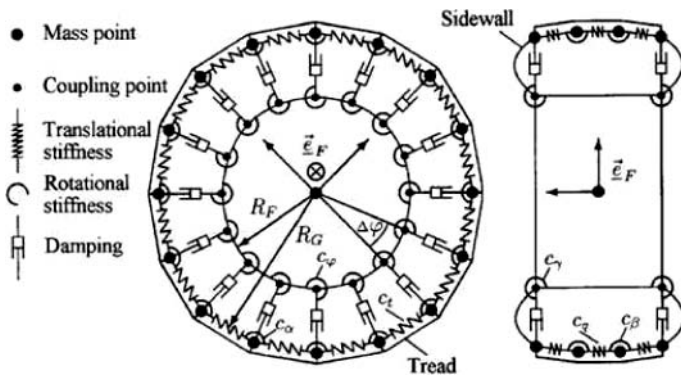
**Figure 50.** Starting situation of a lane change manoeuvre animation.

Even for a MBS vehicle model a model adaption or expansion may be necessary and useful for special problems. For the simulation of the ride on a rough test road a digitized surface needs to be created out of measurements, see Fig. 51. The comparison of the spectral densities to common spectra, [12], confirm the very bad surface quality of this

test track. To simulate the run of an off road vehicle on this track, e.g. for information about the loading of the suspension systems, a proper tyre model has to be integrate in the system model. A tyre model like shown in Fig. 52, [24], or other complex tyre models, see chapter 2, need to be used to be able to reproduce the influences of local surface structures.

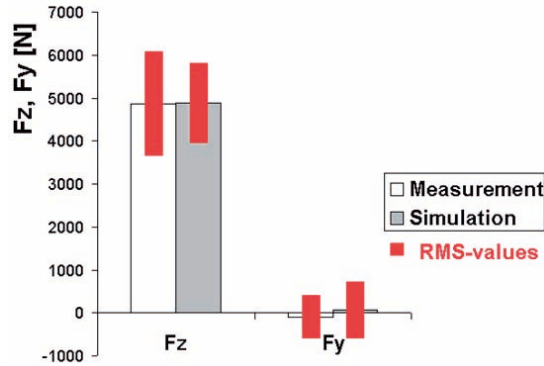


**Figure 51.** Digitized test track surface and corresponding spectral density compared to other surfaces.



**Figure 52.** Structure of complex tyre model, [24], for local surface contact.

Comparing simulation results with measurements, Fig. 53, it can be confirmed that the simulated RMS-values of the tyre forces very well reproduce the reality, opening the calculation area for stress fatigue limits of suspension components.



**Figure 53.** Comparison of tyre forces due to a ride over the surface corresponding to Fig. 51.

One further extension of the vehicle modeling comprises the integration of electronic control systems for all kinds of active interventions in the car or truck. As an example Fig. 54 shows the principal strategy of the yaw motion intervention by the ESP (Electronic Stability Program) to keep the vehicle on the desired path. Thereby a brake force is induced on a front or rear wheel to initiate a correcting yaw moment  $M_{res}$ .

The control system behind the correction activation of the brake force needs on board sensors with respect to individual wheel angular velocities, velocity and yaw velocity of the car, lateral acceleration, steering input. Generally the side slip angle of the car will be estimated by an observer. Adaptations in the hardware must provide the possibility to apply additional individual brake pressures at the wheels for the desired brake forces. And all this has to be tuned corresponding to the special car and other existing systems like an ABS, see e.g. [25].

A further extension of the stability systems is not only applying corrections by braking forces but also by a steering input to the front and/or rear wheels. Fig. 55 wants to give an overall impression of such a closed loop control where additional front or rear steering angles  $\delta_F^A$ ,  $\delta_R^A$  and brake pressure adaptations  $p_i^A$  can be provided, for details look at [26].

The consequences of different control strategies can be noticed for the braking during cornering manoeuvre, Fig. 56. The desired behaviour is to keep the car despite longitudinal deceleration on the same constant radius, what is checked by the corresponding yaw rate and lateral acceleration after the step brake force application. As can be noticed the uncontrolled vehicle (passenger car) shows remarkable deviations in the area of about a deceleration of  $4\text{m/s}^2$ . The comparison with the initial values  $\dot{\psi}_0$ ,  $a_{y0}$  indicate that the car turns to the inner side of the curve with stronger yaw velocity. All 4 control strategies –  $\Delta M_{Y_{aw}}$  stands for a differential braking – can improve this negative behaviour and show increasing differences for very high longitudinal decelerations only.

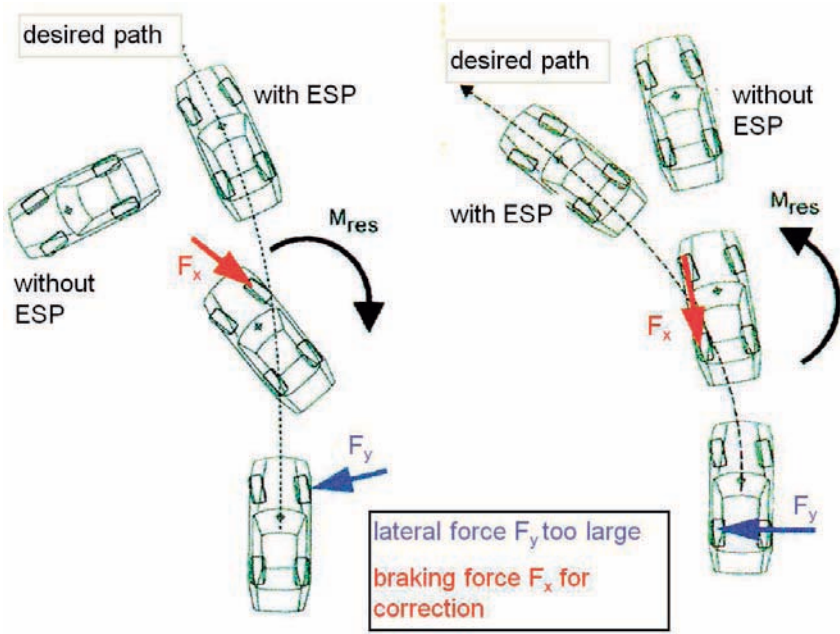


Figure 54. Principal operation of the ESP-system.

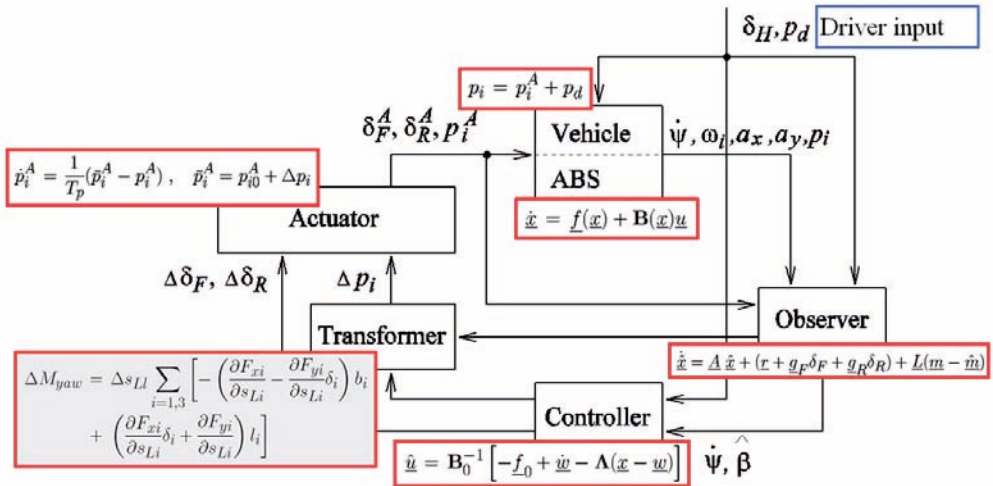
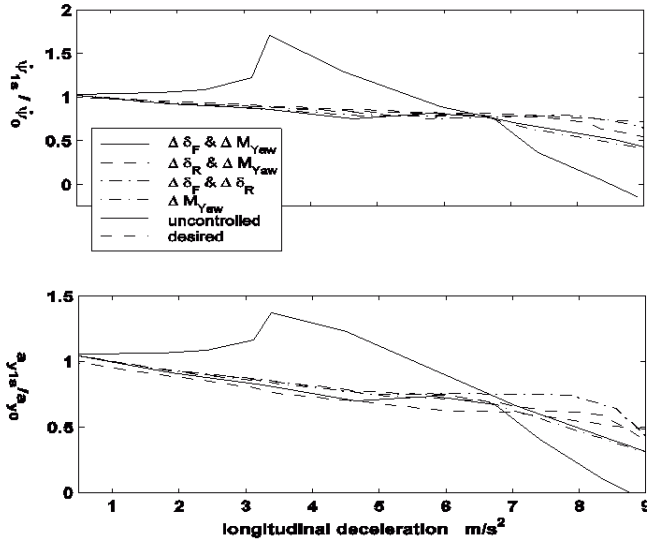


Figure 55. Lateral stability control with brake pressure adaption and additional steering.





**Figure 56.** Braking during cornering: influence of control strategies. Normalized yaw rate and lateral acceleration 1s after braking step input.

A very interesting topic with increasing importance is the completion of the vehicle (and control) model with a driver model. Only then the simulation of closed loop manoeuvres can be used for the design concepts of vehicles. Basically, as e.g. shown in [12], [28], a 2-level driver model is introduced with an anticipatory feed forward control and a compensatory closed loop that provides the steer angle corrections  $\Delta\delta_c$  by an estimation of the future vehicle deviation to the desired path. So the feed forward control, based on the driver perception (at time  $t$ ) of the oncoming road curvature  $\kappa(t + T_i)$  provides the main steering input  $\delta_S$  and can also include some kind of trajectory learning.

In Fig. 57 a modified 2-level driver model is presented, that shows these characteristic features, [27]. Moreover by the explicit input of the driving velocity  $v_c$  a "longitudinal" driver model can be added.

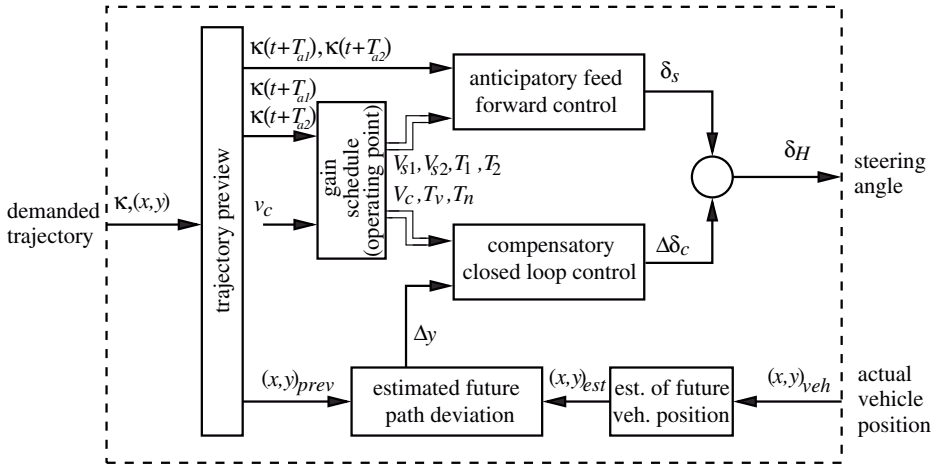
Corresponding to Fig. 57 for the anticipatory control the driver behaviour at time  $t$  is described using 2 predicted curvatures  $\kappa$  for the considered time preview  $T_{a1}, T_{a2}$ .

$$T_2^2 \ddot{\delta}_s(t) + T_1 \dot{\delta}_s(t) + \delta_s(t) = V_{s1} \kappa(t + T_{a1}) + V_{s2} \kappa(t + T_{a2}) \quad (5.1)$$

The parameters of this equation (5.1) are determined by a comparison to the linear (or locally linearized) 2-wheel model based on the interpretation that by experience the driver knows the characteristics of his car.

For the compensatory control loop the driver reaction to the estimated future path deviation  $\Delta y$  can be presented by the transfer function

$$G_C(s) = \frac{\Delta\delta_c(s)}{\Delta y(s)} = V_c \frac{1 + T_v s}{1 + T_n s} e^{-T_r s} \quad (5.2)$$



**Figure 57.** Two level driver model for lateral dynamics.

with the reaction time delay  $T_r$ . The other parameters are to be determined via the control design and the hereby used specifications.

The estimation of the path deviation first uses the transfer function  $G_V$  of the linear (or linearized) vehicle model – here described symbolically by the constants  $K_i$  – and a Taylor series expansion with the preview time  $T_a$ :

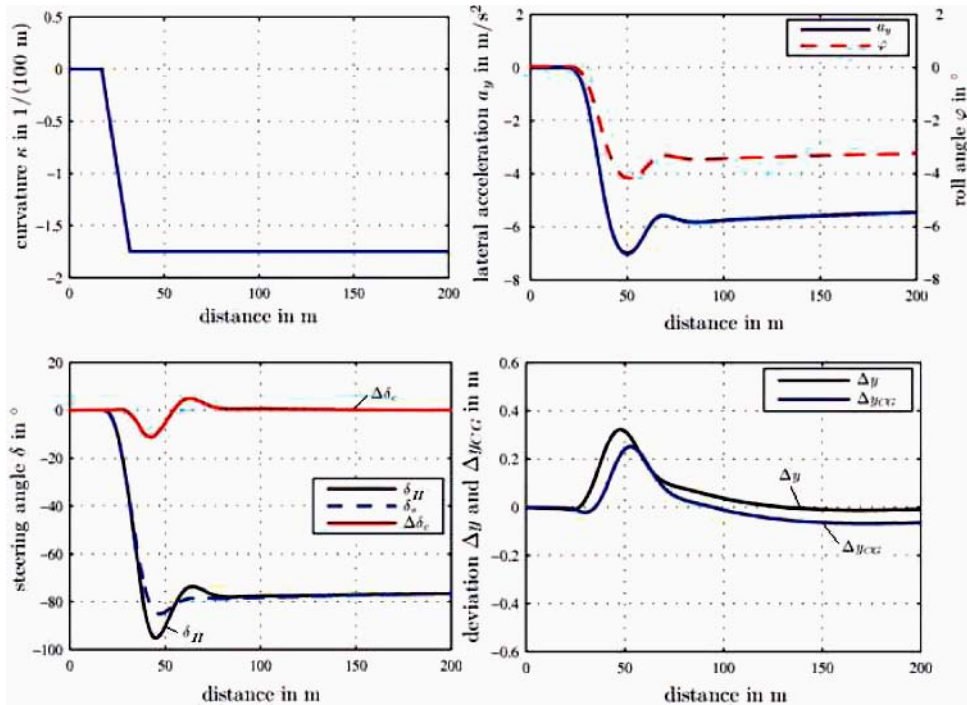
$$G_V(s) = \frac{\Delta y_{CG}(s)}{\Delta \delta_c(s)} = \frac{v_c (K_3 s^2 + K_4 s + K_5)}{i_S s^2 (s^2 + K_1 s + K_2)}, \tag{5.3}$$

$$G_P(s) = \frac{\Delta y(s)}{\Delta y_{CG}(s)} = 1 + T_a s + \frac{T_a^2}{2} s^2. \tag{5.4}$$

In Fig. 58 the vehicle-driver reaction to a given curvature of an entry into a curve is presented. There is no real problem for the driver to enter the curve despite the lateral acceleration indicates a challenging manoeuvre with higher final acceleration than usual for regular driving. The deviation  $\Delta y$  shows that the large values occur just after the end of the entry clothoide. Small deviations remain in the nearly steady state area after a distance of 100m. Remarkable is the dominant part of the anticipatory steering wheel angle  $\delta_s$  for the whole steering wheel angle  $\delta_H$ .

The features and future aspects of an advanced vehicle system model may be emphasized by some keywords:

- Based on simple models: "hand made" 3D-models, 2- and 4-wheel models for control design
- Detailed modeling of vehicle components and their integration within MBS software
- Road surface as a system component together with proper tyre models
- Control devices as essential components of the system "vehicle"



**Figure 58.** Simulation of driver behaviour for a given curve entry using an MBS-vehicle model; velocity  $v_c \approx 19$  m/s.

- Driver assistance or driver reaction replacement in emergency situations by control systems.
- Driver modeling for closed loop manoeuvres
- Increased use of simulation together with optimization for the vehicle development and design

Despite or better because of the availability and relative easy implementation of vehicle-control-driver-systems you should be aware of the necessary balance between modeling details, knowledge of parameters and interpretation of results.

## Bibliography

- [1] Gillespie T.D.: *Fundamentals of Vehicle Dynamics*, Society of Automotive Engineers, Inc., Warrendale, PA, USA, 1992.
- [2] Lugner P., Plöchl M. (editors): *Tyre Models for Vehicle Dynamics Analysis*, Vehicle System Dynamics (Supplement) 43, 2005.
- [3] Pacejka H.: *Tyre and vehicle dynamics*, Butterworth-Heinemann, 2002.
- [4] Pacejka H., Bakker E.: *The Magic Formula Tyre Model*, Proceedings of the 1st International Colloquium on Tyre Models for Vehicle Dynamics Analysis, Swets and Zeitlinger, 1993.
- [5] Lugner P., Mittermayr P.: *A measurement based tyre characteristics approximation*, Proceedings of the 1st International Colloquium on Tyre Models for Vehicle Dynamics Analysis, Swets and Zeitlinger, 1993.
- [6] Segel L.: *The Tyre as a Vehicle Component in Mechanics of Transportation Systems*, ASME AMD, Vol. 15, 1975.
- [7] Wiegner P.: *Über den Einfluß von Blockverhindern auf das Fahrverhalten von Personenkraftwagen bei Panikbremsung*, Dissertation (im Anhang), TU Carolo-Willhelmina, Braunschweig, 1973.
- [8] Kortüm W., Lugner P.: *Systemdynamik und Regelung von Fahrzeugen*, Springer Verlag, 1994.
- [9] Gipser M.: *FTire: a physically based application-oriented tyre model for use with detailed MBS and finite-element suspension models*, Vehicle System Dynamics (Supplement) 43, 2005.
- [10] Lugner P., Pacejka H., Plöchl M.: *Recent advances in tyre models and test procedures*, Vehicle System Dynamics 43, 2005.
- [11] Genta G.: *Motor Vehicle Dynamics, Modelling and Simulation*, World Scientific Publishin Co., 1997.
- [12] Mitschke M., Wallentawitz H.: *Dynamik der Kraftfahrzeuge*, Springer Verlag, 4. Auflage, 2004.
- [13] Hucho W.-H.: *Aerodynamik des Automobils*, Springer Verlag, 3. Auflage, 1999.
- [14] Zomotor A.: *Fahrwerktechnik: Fahrverhalten*, Vogel Vuchverlag Würzburg, 1987.
- [15] Richter, B.: *Schwerpunkt der Fahrzeugdynamik*, Verlag TÜV Rheinland, 1990.
- [16] Popp K., Schiehlen W.: *Fahrzeugdynamik*, Teubner Verlag Stuttgart, 1993.
- [17] Lugner P.: *Horizontal Motion of Automobiles*, CISM Course and Lectures No. 274 edited by W.O. Schiehlen, Springer Verlag, 1982.
- [18] *Bosch Automotive Handbook*, 6th edition, Robert Bosch GmbH Distribution, SAE, USA, 2004.
- [19] Lugner P., Plöchl M.: *Dynamic traction limitations of a passenger car with 4-wheel drive*, Int. J. Vehicle Design, Vol. 25, No. 3, 2001.
- [20] Cebon D.: *Handbook of Vehicle Road Interaction*, Swets and Zeitlinger, 1999
- [21] Kopetz Ch., Lugner P., Rutz R.: *Model based Vehicle Dynamic Analysis*, ATZ worldwide, 11/2004. Modellbasierte Fahrdynamikanalyse, ATZ 11/2004, p. 1030-1037.
- [22] Ammon D.: *Modellbildung und Systementwicklung in der Fahrzeugdynamik*, Teubner Verlag Stuttgart, 1997.

- [23] Schiehlen W. (editor): *Multibody Sstems Handbook*, Springer Verlag, 1990.
- [24] Oertel Ch., Eichler M., Faudre A.: *RMOD-K, Version 5.2, Manual*, CA Entwicklung, gedas GmbH, 1998.
- [25] van Zanten A., Erhardt R., Pfaff G.: *FDR-Die Fahrdynamikregelung von Bosch*, ATZ 11/1994, p. 674-689.
- [26] Heinzl Ph., Lugner P., Plöchl M.: *Stability Control of a Passenger Car by Combined Additional Steering and Unilateral Braking*, Vehicle System Dynamics (Supplement) 37, 2003.
- [27] Edelmann J., Plöchl M., Reinalter W., Tieber W.: *A Passenger Car Driver Model for High Lateral Accelerations*, Vehicle System Dynamics, in press.
- [28] MacAdam CC.: *Understanding and Modeling the Human Driver*, Vehicle System Dynamics 40, 2003.

# Dynamics of Motorcycles: Stability and Control

Robin S. Sharp

Imperial College London, London SW7 2AZ, UK

**Abstract.** The discussion presumes that theory aligns closely with practice in respect of major issues. References in support are quoted. A relatively simple historic mathematical model is used to establish symbolic equations of motion and to extract some insight from them. Numerical methods are then employed with a succession of elaborations of the simple model, to establish what level of model is necessary to represent what behaviour. Eigenvalues, eigenvectors and frequency responses are shown. To predict a full range of small-perturbation properties well, tyre sideslip has to be allowed and relaxation delays, frame torsional compliance and rider-upper-body freedom have to be included. More involved and accurate modelling is described. A fully detailed model is linearised for small perturbations from straight running and linear optimal preview control theory is applied to finding good rider steering control strategies. Those strategies are applied to path tracking simulations. Conclusions are drawn at the end.

## 1 Introduction

Road vehicle handling problems involve interactions between man and machine. A full treatment of the subject requires consideration of both human factors and vehicle engineering. Human factors in motorcycle dynamics include the rider as a structural component of the total system and as a controlling element, sometimes being required to stabilize the uncontrolled machine and always being required to control the path followed. Human aspects of the full problem are much more complex than the engineering ones and it is necessary to deal with them in a less rigorous and exact manner, due to obvious difficulties of experimentation and analysis.

Relating to handling dynamics, bicycles and motorcycles have much in common, although parametric descriptions and speed ranges are quite different for the two cases. Attention is paid here to the motorcycle problem. Corresponding bicycle material can be found in Limebeer and Sharp (2006). It has been established that state-of-the-art models mimic the main behaviour of single-track vehicles, so that an understanding of the subject can be obtained through study of mathematical models. Such models can be built at many different levels of detail, the simplest of them exposing the basic physics involved but being limited in terms of fidelity. Engagement with quite elaborate models, built with the aid of computer-aided-engineering software, is essential to appreciation of the full spectrum of motorcycle stability and response properties.

In the following sections, the free-control motorcycle is studied at a relatively simple level from a symbolic viewpoint. We concentrate on the free-control case, since it has

been demonstrated that the fixed-control motorcycle is divergently unstable at all but high speeds (Sharp, 1971), making it certain that single-track vehicles are operated in free-control almost exclusively. The accent is on simplicity and generality. When the limits of the symbolic approach have been reached, numerical results are relied on. We step through a sequence of modelling levels, examining the changes in the predicted stability properties as features are added. The influences of modelling assumptions on the behaviour are revealed and conclusions regarding modelling needs are reached. Advanced models and the relationship between predicted and measured behaviour are discussed. The role of the rider in tracking a desired path is considered and analysed from the point of view of linear optimal preview control theory.

## 2 Basic Theory of Motorcycle Stability and Control

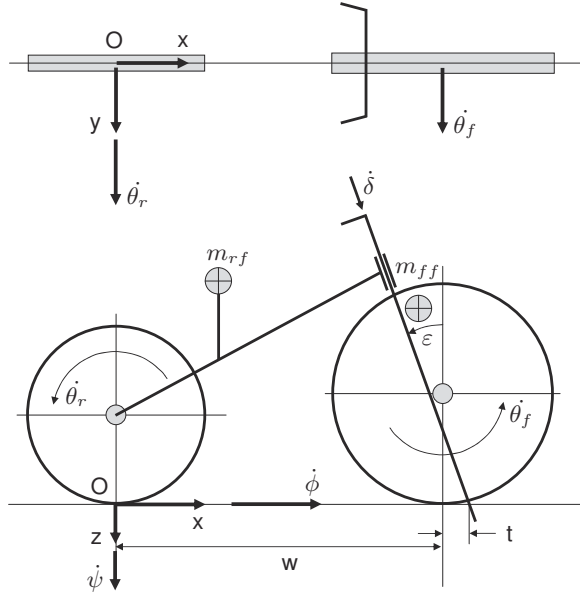
Whipple (1899) published a mathematical model of a bicycle which still serves a useful purpose in illuminating the mechanics of single-track vehicles. Whipple's model is shown diagrammatically in Figure 1. The model consists of two frames joined by a revolute joint at the steering hinge, with each frame supporting an axi-symmetric wheel that is free to spin. Roll and steer freedoms are allowed, while the forward speed is presumed constant. For stability analysis, the steering system is free to steer itself. Whipple apparently had the wisdom to see that this self-steering action is vital to the functioning of a single-track vehicle, a feature that distinguishes the dynamics of motorcycles strongly from the dynamics of cars. The rider is considered a rigid extension of the main frame, dominating its mass properties in the bicycle case. The front frame also has distributed mass. To reduce the parametric description of the motorcycle to its lowest level of complexity, the mass properties of each axi-symmetric road wheel are included with the appropriate frame, except that the spin inertia needs to be specified separately in each case.

The tyres in the Whipple model are treated as providing pure-rolling, nonholonomic constraints on the longitudinal, lateral and yawing motions of the vehicle. The equations of motion of the system established by Whipple, for small perturbations from straight-running, have been thoroughly checked for accuracy (Schwab et al., 2004, 2005b,a) and expressed in an economical form through the use of intermediate variables, shortly to be specified. A minimal set of symbols for describing a machine is given in Table 1, in which the parameter values come from Sharp (1971), reduced to the form given in Schwab et al. (2004, 2005b,a). In particular, the front frame inertias are referred to the same xyz-directions as are used for the rear frame, and the wheel masses and in-plane inertias are lumped together with the appropriate frame properties. Also, we use AutoSim's sign convention for products of inertia (Mitiguy and Reckdahl, 1998), for example,  $I_{xz} = - \int xz \cdot dm$  in which  $m$  denotes mass. Whipple's model and the parameter values specified in Table 1 constitute the basic motorcycle.

### 2.1 Symbolic analysis

Following Schwab et al. (2004, 2005b,a), with the additional economy that comes from lumping the wheels with the frames, we use intermediate variables as follows.

$$m_t = m_{rf} + m_{ff} \tag{2.1}$$



**Figure 1.** Diagrammatic motorcycle with front frame mass,  $m_{ff}$ , and rear frame mass,  $m_{rf}$ , adapted from Schwab et al. (2004, 2005b,a). Contrasting with the references, front and rear wheel masses and inertias are incorporated into the respective frames, as appropriate for axially symmetric bodies. Only the spin inertia of each wheel needs to be specified individually. The number of parameters needed to define the motorcycle is thereby reduced to a minimum.

$$x_t = (x_{rf}m_{rf} + x_{ff}m_{ff}) / m_t \tag{2.2}$$

$$z_t = (z_{rf}m_{rf} + z_{ff}m_{ff}) / m_t \tag{2.3}$$

$$T_{xx} = B_{xx} + C_{xx} + m_{rf}z_{rf}^2 + m_{ff}z_{ff}^2 \tag{2.4}$$

$$T_{xz} = B_{xz} + C_{xz} - m_{rf}x_{rf}z_{rf} - m_{ff}x_{ff}z_{ff} \tag{2.5}$$

$$T_{zz} = B_{zz} + C_{zz} + m_{rf}x_{rf}^2 + m_{ff}x_{ff}^2 \tag{2.6}$$

$$u = (x_{ff} - w - t) \cos \epsilon - z_{ff} \sin \epsilon \tag{2.7}$$

$$F_{\epsilon\epsilon} = m_{ff}u^2 + C_{xx}\sin^2 \epsilon + 2C_{xz} \sin \epsilon \cos \epsilon + C_{zz}\cos^2 \epsilon \tag{2.8}$$



**Table 1.** Symbols and parameter values of the basic motorcycle. Note that wheel masses and in-plane inertias are lumped in with the relevant parent bodies. Also, although pitch inertias  $B_{yy}$  and  $C_{yy}$  are given, these symbols do not appear in the linear equations of motion for the system and the values do not affect the small-perturbation-from-straight-running behaviour. A particular laterally-symmetric motorcycle at this small-perturbation level is defined by 19 parameters. Acceleration due to gravity has the standard value  $9.81 \text{ m/s}^2$ .

Parameter	Symbol	Value
Wheelbase	$w$	1.4145 m
Trail	$t$	0.13 m
Steer axis inclination	$\epsilon$	0.4715 rad
Forward speed	$V$	variable m/s
Rear wheel		
Radius	$R_{rw}$	0.3048 m
Mass	$m_{rw}$	0 kg
Moments of inertia	$(0, A_{yy}, 0)$	$(0, 1.0508, 0) \text{ kg-m}^2$
Rear frame		
Mass centre position	$(x_{rf}, y_{rf}, z_{rf})$	$(0.4798, 0.0, -0.6157) \text{ m}$
Mass	$m_{rf}$	217.45 kg
Moments of inertia	$\begin{bmatrix} B_{xx} & 0 & B_{xz} \\ & B_{yy} & 0 \\ sym & & B_{zz} \end{bmatrix}$	$\begin{bmatrix} 31.184 & 0 & -1.7354 \\ & 45.0 & 0 \\ & & 21.069 \end{bmatrix} \text{ kg-m}^2$
Front frame		
Mass centre position	$(x_{ff}, y_{ff}, z_{ff})$	$(1.3335, 0.0, -0.4672) \text{ m}$
Mass	$m_{ff}$	30.467 kg
Moments of inertia	$\begin{bmatrix} C_{xx} & 0 & C_{xz} \\ & C_{yy} & 0 \\ sym & & C_{zz} \end{bmatrix}$	$\begin{bmatrix} 1.0714 & 0 & -0.3207 \\ & 1.3 & 0 \\ & & 0.6059 \end{bmatrix} \text{ kg-m}^2$
Front wheel		
Radius	$R_{fw}$	0.3048 m
Mass	$m_{fw}$	0 kg
Moments of inertia	$(0, D_{yy}, 0)$	$(0, 0.7186, 0) \text{ kg-m}^2$

$$F_{\epsilon x} = -m_{ff}uz_{ff} + C_{xx} \sin \epsilon + C_{xz} \cos \epsilon \quad (2.9)$$

$$F_{\epsilon z} = m_{ff}ux_{ff} + C_{xz} \sin \epsilon + C_{zz} \cos \epsilon \quad (2.10)$$

$$f = t \cos \epsilon / w \quad (2.11)$$

$$S_f = D_{yy} / R_{fw} \quad (2.12)$$

$$S_t = A_{yy} / R_{rw} + D_{yy} / R_{fw} \quad (2.13)$$

$$S_u = m_{ff}u + fm_t x_t. \quad (2.14)$$

The small-perturbation linear equations of motion, written in the brief form enabled by the use of the intermediate variables appearing in Equations (2.1) to (2.14), are:

$$\begin{aligned} T_{xx}\ddot{\phi} + gm_t z_t \phi = & -(F_{\epsilon x} + fT_{xz})\ddot{\delta} + \\ (fm_t z_t - fS_t - S_f c_\epsilon - T_{xz} c_\epsilon / w) V \dot{\delta} + & \\ \{gS_u - (S_t - m_t z_t) c_\epsilon V^2 / w\} \delta & \end{aligned} \quad (2.15)$$

and

$$\begin{aligned} (F_{\epsilon \epsilon} + 2fF_{\epsilon z} + f^2 T_{zz}) \ddot{\delta} + \{F_{\epsilon z} c_\epsilon / w + f(S_u + T_{zz} c_\epsilon / w)\} V \dot{\delta} + & \\ \{(S_u + S_f s_\epsilon) c_\epsilon V^2 / w - gS_u s_\epsilon\} \delta = & -(F_{\epsilon x} + fT_{xz}) \ddot{\phi} + \\ (fS_t + S_f c_\epsilon) V \dot{\phi} + gS_u \phi + M_\delta & \end{aligned} \quad (2.16)$$

These equations can be written in terms of the original motorcycle parameters (with the aid of the MATLAB symbolic toolbox) and, for compactness, the substitutions:  $u$ , the perpendicular distance from the front frame mass centre back to the steer axis, from (2.7);  $S_f$  from (2.12);  $S_t$  from (2.13);  $\sin \epsilon = s_\epsilon$ ;  $\cos \epsilon = c_\epsilon$ ;  $m_x = m_{rf}x_{rf} + m_{ff}x_{ff}$ ;  $m_z = m_{rf}z_{rf} + m_{ff}z_{ff}$ ;  $m_{xz} = m_{rf}x_{rf}z_{rf} + m_{ff}x_{ff}z_{ff}$ ;  $m_{x2} = m_{rf}x_{rf}^2 + m_{ff}x_{ff}^2$ ;  $m_{z2} = m_{rf}z_{rf}^2 + m_{ff}z_{ff}^2$ ; and  $t_n = t/w$ ; yielding

$$\begin{aligned} (B_{xx} + C_{xx} + m_{z2}) \ddot{\phi} + m_z g \phi = & \{m_{ff}uz_{ff} - C_{xx} s_\epsilon - C_{xz} c_\epsilon + \\ t_n c_\epsilon (m_{xz} - B_{xz} - C_{xz})\} \ddot{\delta} - \{t_n S_t + S_f + (B_{xz} + C_{xz} - m_{xz}) / w - & \\ m_z t_n\} c_\epsilon V \dot{\delta} + \{g(m_{ff}u + m_x t_n c_\epsilon) - (S_t - m_z) c_\epsilon V^2 / w\} \delta & \end{aligned} \quad (2.17)$$

and

$$\begin{aligned}
& \{m_{ff}u^2 + C_{xx}s_\epsilon^2 + 2C_{xz}s_\epsilon c_\epsilon + C_{zz}c_\epsilon^2 + 2t_n c_\epsilon (m_{ff}ux_{ff} + C_{xz}s_\epsilon + C_{zz}c_\epsilon) + \\
& (t_n c_\epsilon)^2 (B_{zz} + C_{zz} + m_{x2})\} \ddot{\delta} + \{(m_{ff}ux_{ff} + C_{xz}s_\epsilon + C_{zz}c_\epsilon) c_\epsilon/w + t_n c_\epsilon \\
& [m_{ff}u + m_x t_n c_\epsilon + (B_{zz} + C_{zz} + m_{x2}) c_\epsilon/w]\} V \dot{\delta} + \{-g (m_{ff}u + m_x t_n c_\epsilon) s_\epsilon \\
& + (m_{ff}u + m_x t_n c_\epsilon + S_f s_\epsilon) c_\epsilon V^2/w\} \delta = \{m_{ff}uz_{ff} - C_{xx}s_\epsilon - C_{xz}c_\epsilon + \\
& t_n c_\epsilon (m_{xz} - B_{xz} - C_{xz})\} \ddot{\phi} + \{t_n S_t + S_f\} c_\epsilon V \dot{\phi} + \\
& g (m_{ff}u + m_x t_n c_\epsilon) \phi + M_\delta \quad (2.18)
\end{aligned}$$

as the equations of motion of the basic motorcycle.

These equations can be interpreted as follows: In (2.17), the left-hand-side terms describe the dynamics of an inverted pendulum with time-to-double given by  $0.693\sqrt{(B_{xx} + C_{xx} + m_{z2})/(m_z g)}$ . The right-hand-side terms can be seen as forcing for the pendulum, determining whether or not it remains under control. The terms are of the form  $\{p_2 \ddot{\delta} + p_1 V \dot{\delta} + (p_{00} + p_{01} V^2) \delta\}$ , constituting steer acceleration, steer velocity and steer angle parts, the last two of which are speed-dependent.

In (2.18), steering system inertia:

$$m_{ff}u^2 + C_{xx}s_\epsilon^2 + 2C_{xz}s_\epsilon c_\epsilon + C_{zz}c_\epsilon^2 + 2t_n c_\epsilon (m_{ff}ux_{ff} + C_{xz}s_\epsilon + C_{zz}c_\epsilon) + (t_n c_\epsilon)^2 (B_{zz} + C_{zz} + m_{x2})$$

steer damping:

$$\{(m_{ff}ux_{ff} + C_{xz}s_\epsilon + C_{zz}c_\epsilon) c_\epsilon/w + t_n c_\epsilon [m_{ff}u + m_x t_n c_\epsilon + (B_{zz} + C_{zz} + m_{x2}) c_\epsilon/w]\} V$$

and steering stiffness:

$$-g (m_{ff}u + m_x t_n c_\epsilon) s_\epsilon + (m_{ff}u + m_x t_n c_\epsilon + S_f s_\epsilon) c_\epsilon V^2/w$$

can be recognised on the left-hand side. On the right side, there are self-steering terms of the form  $(g_2 \ddot{\phi} + g_1 V \dot{\phi} + g_0 \phi)$ . These terms represent roll acceleration to steer torque feedback, roll velocity to steer torque feedback and roll displacement to steer torque feedback influences. With no rider control, ( $M_\delta = 0$ ), these feedback terms, together with the steer inertia, damping and stiffness coefficients, determine the steering motions and, in turn from (2.17), the rolling motions of the machine. The most critical of these activities are those that determine the self-steering and it is instructive to examine these influences in greater detail.

Expressions for the feedback ‘‘gains’’  $g_2$ ,  $g_1 V$  and  $g_0$  come directly from (2.18) as:

$$\begin{aligned}
g_2 &= m_{ff}uz_{ff} - C_{xx}s_\epsilon - C_{xz}c_\epsilon + t_n c_\epsilon (m_{xz} - B_{xz} - C_{xz}) \\
g_1 V &= (t_n S_t + S_f) c_\epsilon V \\
g_0 &= g (m_{ff}u + m_x t_n c_\epsilon)
\end{aligned}$$

Contributions to the self-steering action coming from  $g_2$  derive from the front frame mass distribution, the trail and steering inclination and the rear frame product of inertia. These terms influence higher frequency behaviour.  $g_1 V$  involves rear and front wheel gyroscopic torques proportional to speed, trail and inclination and influences mid-frequency behaviour.  $g_0$  relates to the gravitational torque on the steering system deriving from front frame mass centre offset from the steer axis, mechanical trail and steering

inclination. This  $g_0$  term is influential mainly at low frequencies. The high frequency influence of the front frame mass distribution suggests that this design area is likely to be important in relation to wobble oscillations (Juden, 1987; Limebeer and Sharp, 2006), especially in bicycles, where load may be carried in a front-frame-mounted basket and the load may well be a significant proportion of the frame mass. The various expressions simplify remarkably if the trail is zero.

## 2.2 Numerical analysis

The analysis simplifies but becomes less general if it is put into numerical form. Also, numerical methods can be applied irrespective of the complexity of the plant model employed. In the first instance, we can use the parameter values for the basic motorcycle to develop further understanding of the low-speed balancing problem. Then, continuing to deal only with a linear motorcycle model, root locus analysis will allow examination of stability properties in particular, and of how they vary with assumptions made and with machine parameters. Other standard linear system computations will be useful in assessing the behavioural consequences of operating conditions and design decisions.

**The basic motorcycle.** Employing the parameter values of the basic motorcycle as given in Table 1, the equations of motion for autonomous operation are:

$$121.34\ddot{\phi} - 1453\dot{\phi} = -7.191\ddot{\delta} - 65.822V\dot{\delta} + (123.678 - 96.945V^2)\delta \quad (2.19)$$

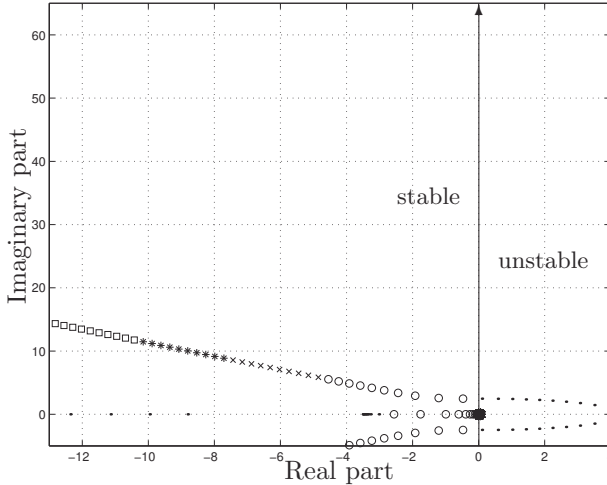
and

$$1.530\ddot{\delta} + 8.394V\dot{\delta} + (8.615V^2 - 56.177)\delta = -7.191\ddot{\phi} + 2.576V\dot{\phi} + 123.68\phi \quad (2.20)$$

The time constant of the inverted pendulum is 0.289 s and the steer angle forcing term is positive for low speeds and negative for higher ones, for which it has the dominant influence on the roll angle of the motorcycle. This forcing term changes sign through 1.13 m/s, so that, in this neighbourhood, the steering velocity forcing term is especially significant. The steering stiffness is negative for low speeds and positive for high ones, changing sign at 2.554 m/s, and the roll angle feedback term is much larger than the velocity and acceleration terms, with respect to the low-speed balancing problem to which the basic motorcycle model principally applies. It will be shown below that the pure-rolling tyre assumption limits the applicability of the model to speeds under about 10 m/s. Thus it is inappropriate to use these numbers in connection with higher speed running.

Now we compute the basic motorcycle's eigenvalues through speed, for small perturbations from a straight-running trim state, Figure 2. At very low speeds, the main features of the basic motorcycle are two divergent modes with times to double 0.1 s and 0.2 s, the former being associated with a ratio of steer to roll of -19, the latter with a ratio of -1, given by the eigenvectors. The more rapidly divergent mode, dominated by steering, is really only of academic interest, since the real tyre, having extended contact with the ground, would develop forces and moments in response to the turnslip occurring

in the modal motion (Pacejka, 2002) which are not represented at all in the basic model, or in any of the other models to be considered. With increasing speed, the two real eigenvalues involved converge on each other and then coalesce at about  $1\text{ m/s}$  to form an oscillatory mode. The mode stabilises at about  $6\text{ m/s}$  and its damping increases with speed thereafter. The basic motorcycle is self-stabilising in the speed range 6 to  $13\text{ m/s}$ .

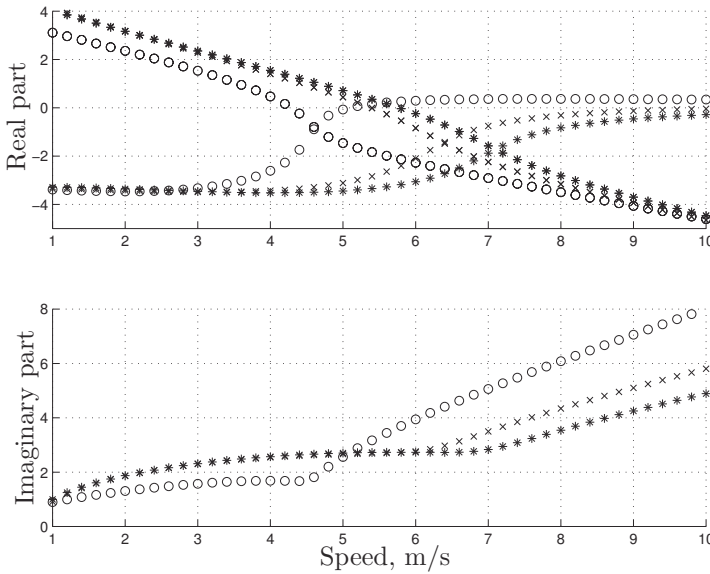


**Figure 2.** Root locus plot for the basic motorcycle over 1 to  $56\text{ m/s}$  speed range. Speed is indicated by the plot symbol, which changes every  $5\text{ m/s}$ . Symbols are used in the order;  $\bullet$ ,  $\circ$ ,  $\times$ ,  $\star$  and  $\square$ . Results for speeds above  $31\text{ m/s}$  lie outside the plot space. The uncontrolled machine will capsize at very low speeds.

**Varying the mechanical trail.** We take the opportunity here to show the influences of trail at only moderate speeds by including cases with zero trail and with twice the basic value of  $0.13\text{ m}$ , Figure 3. For higher speeds, the pure-rolling tyre assumption is inadequate, as will be demonstrated later. Trail influences the stability primarily near to the auto-stable speed range. For increasing trail, this range moves upwards and the worst-speed capsize mode instability is improved. Larger trail values promote easier balancing by the rider in the neighbourhood of  $10\text{ m/s}$  speed. Steering torques required for steady turning increase with trail at low speeds (Sharp, 1971).

**Addition of tyre sideslip.** Let us now elaborate the treatment of the basic motorcycle towards higher fidelity. We include the following features in sequence: Firstly, the tyres are modelled as force producers as opposed to constraining devices, Figure 4. The lateral forces are described as immediate responses to tyre sideslip and camber of the linear form:

$$Y_f = -C_{\alpha_f} \cdot \alpha_f + C_{\gamma_f} \cdot \gamma_f, \quad (2.21)$$



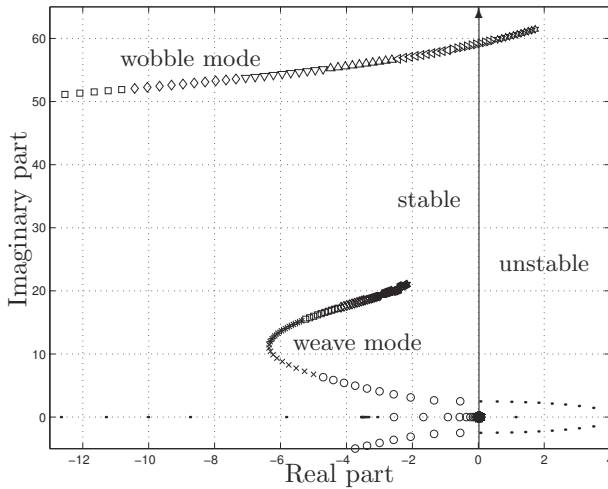
**Figure 3.** Real parts (upper) and imaginary parts (lower) of characteristic roots for the basic motorcycle at modest speeds with variations in trail. The plot symbols; ○, ×, ★ correspond to zero, the standard trail of 0.13 m and twice the standard trail, respectively. At very low speeds, trail is disadvantageous.

in which  $C_{\alpha_f}$  is the front tyre cornering stiffness and  $C_{\gamma_f}$  is the front tyre camber stiffness, with a corresponding equation applying to the rear tyre. Numerical values used come from Sharp (1971) and are:  $C_{\alpha_f} = 11174 \text{ N/rad}$ ;  $C_{\alpha_r} = 938.6 \text{ N/rad}$ ;  $C_{\gamma_f} = 15832 \text{ N/rad}$ ;  $C_{\gamma_r} = 1326.6 \text{ N/rad}$ . In the speed range 0 to 10 m/s the roots are similar to those of the basic motorcycle but above 10 m/s the course of the weave mode damping is radically altered and a wobble mode appears. At high speeds, both weave and wobble appear to be vitally important from a practical point of view. They appear to be potentially problematic, through being insufficiently damped or even unstable.

**Addition of tyre relaxation delays.** Secondly, we include tyre relaxation delays, corresponding physically to the tyre structure lateral compliances (Pacejka, 2002). Tyre lateral forces are now related to their steady-state values by first-order lag equations:

$$\frac{\sigma_f \dot{Y}_f}{U} + Y_f = -C_{\alpha_f} \cdot \alpha_f + C_{\gamma_f} \cdot \gamma_f \tag{2.22}$$

and correspondingly for the rear, where  $\sigma_f$  is the tyre’s relaxation length and  $U$  is the motorcycle speed. We choose relaxation lengths  $\sigma_f = 0.1 \text{ m}$  and  $\sigma_r = 0.12 \text{ m}$ , updated values as compared with those in Sharp (1971), in the light of later information. Root loci are shown in Figure 5. The capsize mode is virtually unchanged, while the weave and



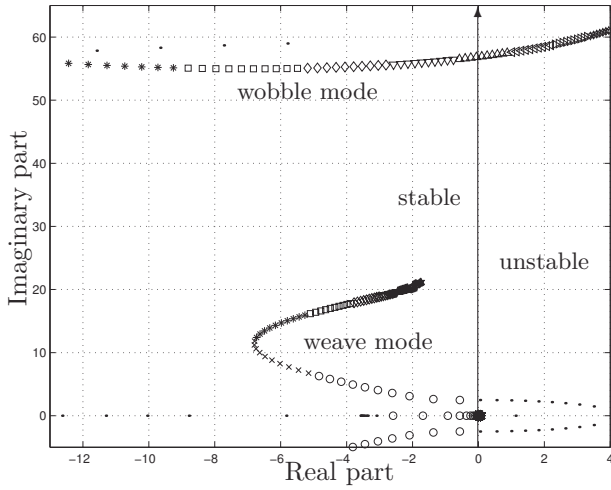
**Figure 4.** Root locus plot for basic motorcycle over 1 to 56  $m/s$  speed range with tyres allowed to sideslip. The tyre lateral force,  $Y_f$ , is given by the expression  $Y_f = -C_{\alpha_f} \cdot \alpha_f + C_{\gamma_f} \cdot \gamma_f$  and correspondingly for the rear. Speed is indicated by the plot symbol, which changes every 5  $m/s$ . Symbols are used in the order;  $\bullet$ ,  $\circ$ ,  $\times$ ,  $\star$ ,  $\square$ ,  $\diamond$ ,  $\nabla$ ,  $\triangle$ ,  $\triangleleft$ ,  $\triangleright$  and hexagon. Numerical values used derive from Sharp (1971) and are:  $C_{\alpha_f} = 11174 N/rad$ ;  $C_{\alpha_r} = 938.6 N/rad$ ;  $C_{\gamma_f} = 15832 N/rad$ ;  $C_{\gamma_r} = 1326.6 N/rad$ . The wobble mode becomes unstable at 49  $m/s$ .

wobble modes are de-stabilised at high speeds, with unstable running of the uncontrolled machine predicted for all speeds above 38  $m/s$ .

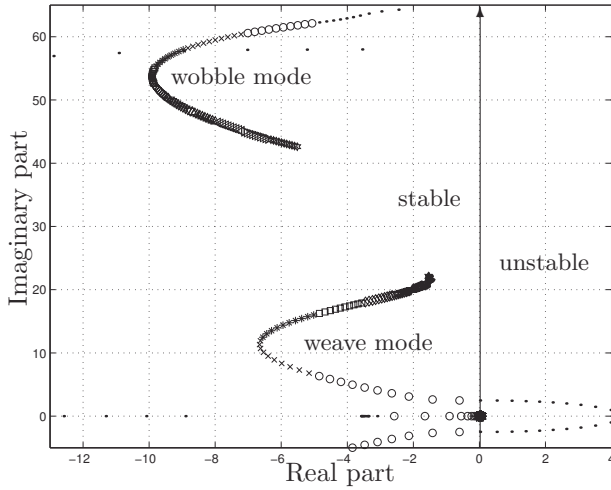
#### Addition of torsional compliance of the main frame at the steering head.

Thirdly, we include torsional compliance in the main frame at the steering head, as found to be important in Eaton (1973); Roe and Thorpe (1976); Sharp and Alstead (1980) and Spierings (1981), Figure 6. The frame stiffness is estimated as 40000  $Nm/rad$  (Clerx, 1977; Koenen and Pacejka, 1982; Giles and Sharp, 1983; Cossalter, 2002) and a modest amount of damping 60  $Nm/(rad/s)$ , to which the behaviour is quite insensitive, is included in parallel with the spring. The compliance modelled represents the contributions of the rear frame and the front forks in series, with lateral fork bending quite possibly being the larger component. The high-speed weave is de-stabilised by the compliance but the high-speed wobble mode is improved by it.

**Addition of rider-upper-body-lean freedom.** Fourthly, a rider-upper-body-roll or -lean freedom is added with parameters as given in Table 2, see Figure 7. Rider data derives from laboratory test results of Nishimi et al. (1985) and represents the decoupled rider as having a natural frequency in roll of about 1.3  $Hz$  with a damping factor of 0.49. This rider freedom has no significant influence on the wobble mode but it splits the weave



**Figure 5.** Root locus plot for basic motorcycle over 1 to 56  $m/s$  speed range with sideslipping tyres and relaxation delays. Plot symbols indicate speed as in Figure 4.  $\sigma_f = 0.1 m$  and  $\sigma_r = 0.12 m$ . The wobble mode now becomes unstable at 38  $m/s$ .



**Figure 6.** Root locus plot for basic motorcycle over 1 to 56  $m/s$  speed range with sideslipping and relaxed tyres and frame torsional compliance. Plot symbols are used to indicate speed as in Figure 4. Numerical values used are  $k_f = 40000 Nm/rad$  and  $c_f = 60 Nm/(rad/s)$ , see Table 2.



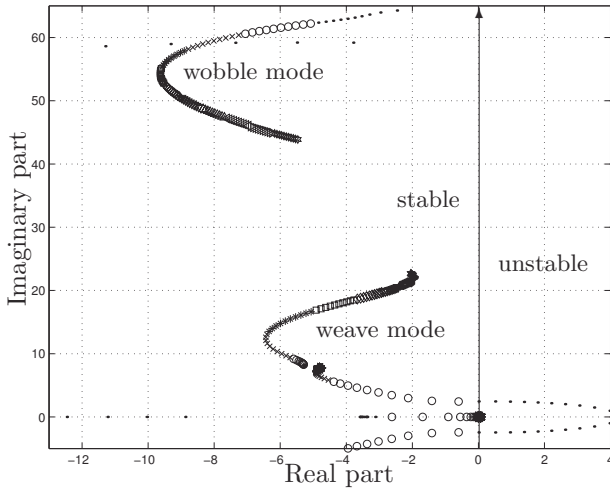
**Table 2.** Those symbols and parameter values of the basic motorcycle which are altered by the inclusion of frame compliance and separation of the rider's upper body from the rear frame. The rider's upper body joint with the rear frame is horizontal at about waist height. Again, the pitch inertias  $B_{yy}$  and  $E_{yy}$  are given but they have no influence on the straight-running behaviour.

Parameter	Symbol	Value
Rear frame		
Mass centre position	$(x_{rf}, y_{rf}, z_{rf})$	(0.4798, 0.0, -0.6157) m
Mass	$m_{rf}$	180 kg
Torsional stiffness	$k_f$	40000 Nm/rad
Torsional damping	$c_f$	60 Nm/(rad/s)
Elastic axis position	$(x_{ef}, y_{ef}, z_{ef})$	(1.2, 0.0, -0.66) m
Moments of inertia	$\begin{bmatrix} B_{xx} & 0 & B_{xz} \\ & B_{yy} & 0 \\ sym & & B_{zz} \end{bmatrix}$	$\begin{bmatrix} 24.716 & 0 & -1.982 \\ & 35.0 & 0 \\ & & 20.246 \end{bmatrix}$ kg-m <sup>2</sup>
Rider upper body		
Mass centre position	$(x_p, y_p, z_p)$	(0.4797, 0.0, -0.95) m
Mass	$m_p$	37.45 kg
Stiffness	$k_p$	350 Nm/rad
Damping coefficient	$c_p$	30 Nm/(rad/s)
Joint height	$h_p$	0.7 m
Moments of inertia	$\begin{bmatrix} E_{xx} & 0 & E_{xz} \\ & E_{yy} & 0 \\ sym & & E_{zz} \end{bmatrix}$	$\begin{bmatrix} 1.4104 & 0 & -0.3207 \\ & 1.6 & 0 \\ & & 0.8236 \end{bmatrix}$ kg-m <sup>2</sup>

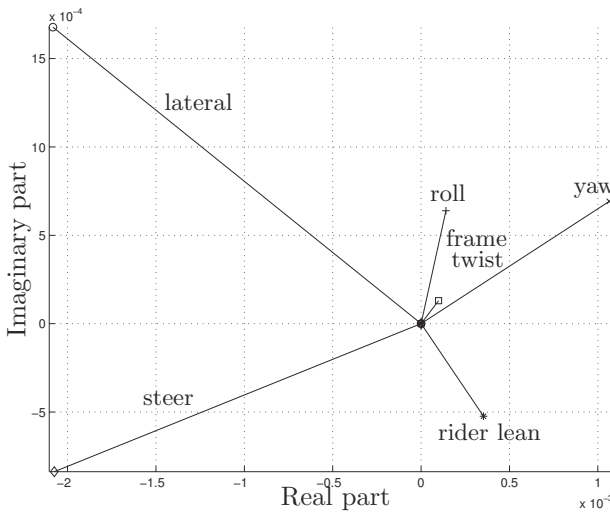
mode into two and stabilises the two weave-type oscillations at the high-speed end of the range. The wobble mode has a circular frequency of 45.0 *rad/s* at 50 *m/s*. Eigenvectors show the participants in the modal motions. These participants illustrate the extent to which the modes of this most elaborate model are influenced by added features like frame compliance and rider-upper-body roll.

**Behaviour of the most detailed of the above motorcycles.** Eigenvectors are shown in Figure 8 for weave, in Figure 9 for rider lean and in Figure 10 for wobble modes. Motion in weave is a complex amalgam of components. Steer and yaw are in phase opposition, as are roll of the main frame and rider roll relative to the main frame. The lateral velocity component is the largest but, of course, this depends on the units used for the diagram. The frame-twist contribution is small, implying that the frame compliance influence on the weave mode will not be so strong. The rider lean mode involves mostly the rider-upper-body-relative-roll and lateral motion of the machine. The main contributors to wobble are steer and frame twist, with steer leading frame twist by about 140° and being in opposition to yaw. The frame compliance is clearly important in determining the nature of this motion and, by implication, its stability.

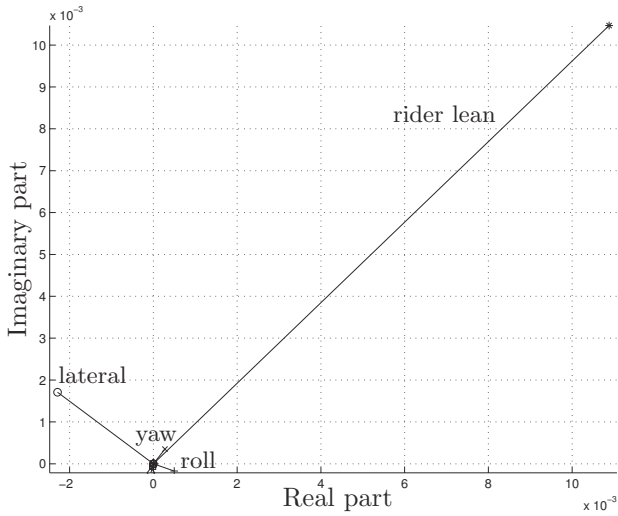
An increasingly common modification to a motorcycle is to add a topbox for luggage



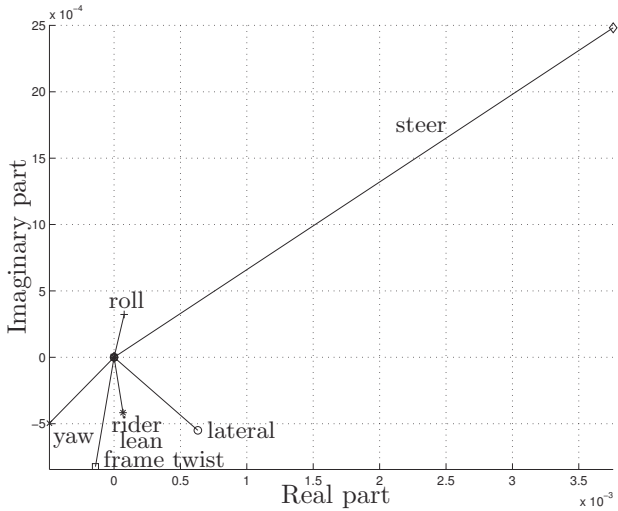
**Figure 7.** Root locus plot for basic motorcycle over 1 to 56  $m/s$  speed range with tyres allowed to sideslip, relaxation delays, frame torsional compliance and rider-upper-body-lean freedom all included. Plot symbols are used to indicate speed as in Figure 4. Numerical values used to describe the rider derive from Nishimi et al. (1985) and are shown in Table 2.



**Figure 8.** Weave mode eigenvector for full system of Figure 7 at 50  $m/s$ . All the variables included are velocities. Lateral velocity is given in  $m/s$ , with the angular velocities expressed in  $rad/s$ . The corresponding eigenvalue is  $-1.98 \pm 22.2i$ .

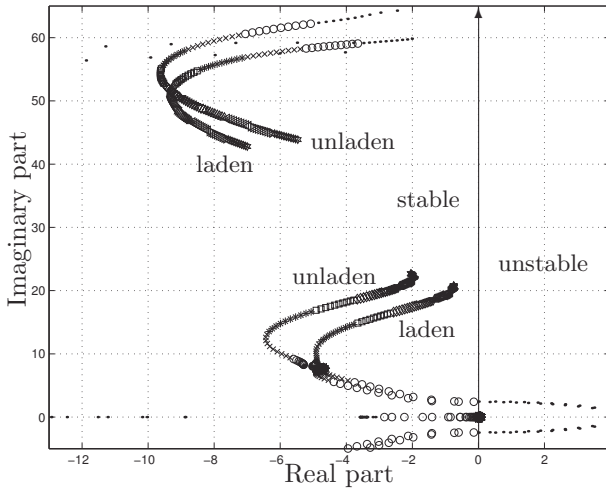


**Figure 9.** Rider lean mode eigenvector for full system of Figure 7 at  $50\text{ m/s}$ . The mode mainly involves rider-upper-body lean relative to the main frame with some lateral velocity phase advanced by about  $90^\circ$ . The corresponding eigenvalue is  $-4.78 \pm 7.75i$ .



**Figure 10.** Wobble mode eigenvector for full system of Figure 7 at  $50\text{ m/s}$ . The mode is dominated by steering, with yaw in phase opposition and significant frame twisting. The corresponding eigenvalue is  $-6.30 \pm 45.0i$ .

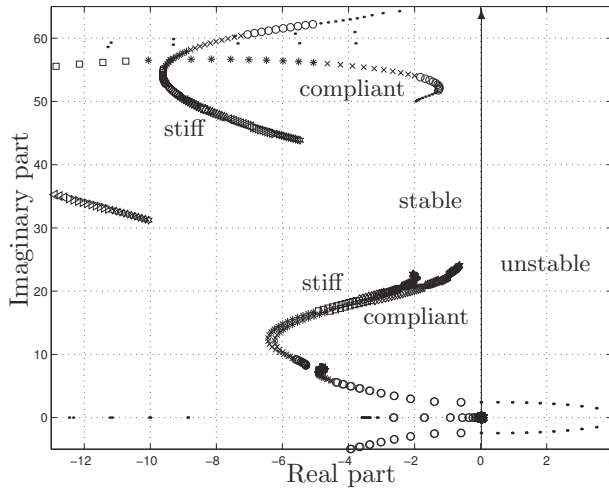
carrying. The influence of such a rigidly attached topbox, with mass  $20\text{ kg}$  and inertias based on the notion that the box contains books and papers, is shown in Figure 11. The additional load destabilises the high-speed weave mode significantly and the practice of specifying a restricted speed range when large rear loads are being carried appears to be entirely justified.



**Figure 11.** Root locus plot for the fully detailed motorcycle of Figure 7 with the addition of a  $20\text{ kg}$  topbox, over  $1$  to  $56\text{ m/s}$  speed range. Plot symbols indicate speed as in Figure 4. Destabilisation of the high-speed weave mode by the rigidly attached rear load is illustrated. At  $50\text{ m/s}$ , the weave mode eigenvalue in the laden condition is  $-0.768 \pm 20.3i$ .

The influences of locking a wheel by braking can be represented approximately by reducing the relevant tyre cornering and camber stiffness coefficients to say one fifteenth of their normal values. Root loci plotted on this basis show that when the front wheel is locked, the motorcycle is oscillatorily unstable at all speeds and when the rear wheel locks, at low speeds there is a very low frequency oscillatory instability and at higher speeds, overdamped modes are unstable. Control of the machine in the first case will be impossible, without unlocking the front wheel, since steering control inputs will leave the tyre force system virtually unaffected. In the second case, steering will be effective but stabilisation will be difficult.

The marked influence of frame compliance is further demonstrated by the root locus plot for the full system with the frame stiffness halved, Figure 12. The manner in which the wobble mode damping varies with speed is radically changed by reducing the stiffness. With a stiff frame the wobble damping increases with speed and then decreases again, predicting wobble problems at high speeds. With a compliant frame, the trend is opposite and gives rise to the prediction of what is well known in practice, that a worst-speed for wobble can exist, at quite low or moderate speeds (Eaton, 1973; Roe and Thorpe, 1976; Dunlop Tyre Company, c1977; Verma et al., 1980; Limebeer et al., 2002).

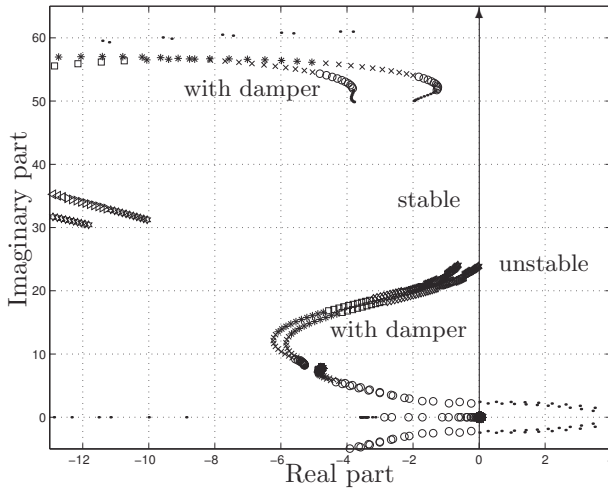


**Figure 12.** Root locus plot for the fully detailed motorcycle of Figure 7 over 1 to 56  $m/s$  speed range, with the corresponding results when the frame stiffness is halved. Plot symbols indicate speed as in Figure 4. The large influences of the frame stiffness in destabilising the weave mode and re-shaping the wobble locus are illustrated.  $k_f = 20000 \text{ Nm/rad}$  and  $c_f = 42.4 \text{ Nm/(rad/s)}$ .

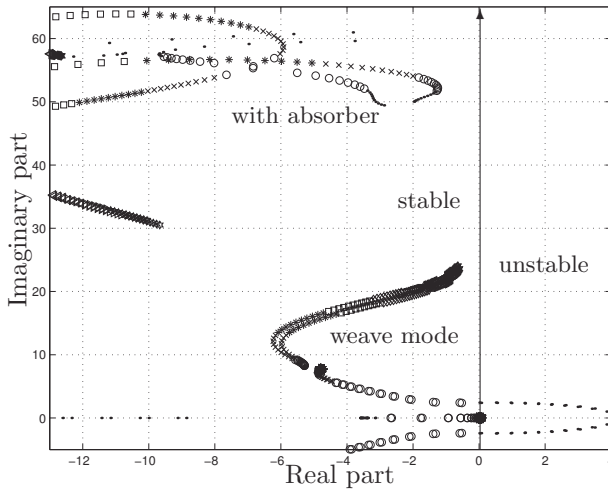
If the wobble mode is insufficiently damped, it can be improved by adding a steering damper (Sharp, 1971, 1978, 1985, 1994, 2001). However, the damper will destabilise the high-speed weave motion and will require the rider to use larger steering control torques, which may not be so comfortable. The stability influences are shown in Figure 13.

Since there is a convenient separation in frequency between wobble and weave, it is apparent that what is needed instead of a steering damper is a device that behaves like a damper for wobble frequencies but not for weave frequencies. One approach to improvement is to add a damped resonant vibration absorber to the steering system, tuned to the frequency of the wobble mode for the speed at which it is minimally damped. Application of this approach to the motorcycle of Figure 12 with reduced frame stiffness, for which the relevant circular frequency is  $52 \text{ rad/s}$ , is illustrated in Figure 14. The dynamic absorber consists of a small flywheel mounted on the rear frame and connected to it through a damper, geared to the rotation of a second shaft with ratio 20, with a torsion spring connecting the second shaft to the handlebars. The flywheel inertia is  $0.003 \text{ kgm}^2$ , the spring rate is  $10.04 \text{ Nm/rad}$  and the damper coefficient is  $0.0707 \text{ Nm/(rad/s)}$ . The vibration absorber resonates at  $57.85 \text{ rad/s}$  circular frequency and is large enough to add significant damping to the wobble mode, while it does not prejudice the weave mode noticeably.

An alternative to the tuned absorber is a mechanical network (Evangelou et al., 2004, 2006a,c) or an active steering compensator (Evangelou et al., 2006b). The mechanical network consists of two parallel shafts mounted in bearings to the rear frame, with one of

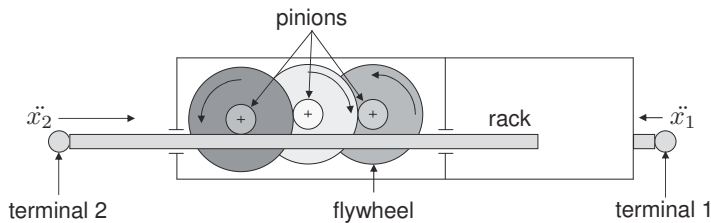


**Figure 13.** Root locus plots for the fully detailed compliant motorcycle of Figure 12 over 1 to 56  $m/s$  speed range with and without a steering damper of coefficient  $4 Nm/(rad/s)$  added. Plot symbols indicate speed as in Figure 4. The influences of the damper in stabilising the wobble and in destabilising the weave are illustrated.



**Figure 14.** Root locus plot for the fully detailed motorcycle of Figure 12, with and without a tuned vibration absorber, over 1 to 56  $m/s$  speed range. The absorber is a rotary geared flywheel subsystem with inertia  $0.003 kgm^2$ , gear ratio 20, stiffness  $10.04 Nm/rad$  and damper coefficient  $0.0707 Nm/(rad/s)$ , as described in the text. Plot symbols indicate speed as in Figure 4. The absorber stabilises the wobble mode without worsening the weave mode.

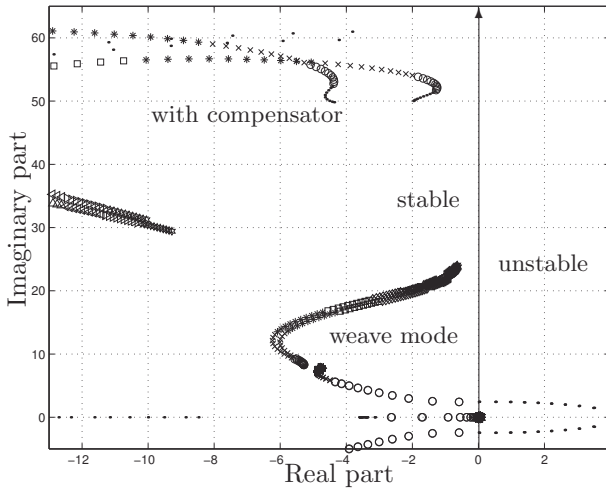
the shafts itself supporting a small flywheel, with spin inertia  $0.000039 \text{ kgm}^2$ , in bearings. A gear train constrains the flywheel to spin with ratio 40 according to the difference between the spin velocities of the two shafts. One shaft is joined to the steering system by a torsion spring of rate  $169 \text{ Nm/rad}$ , while the other shaft is joined to the rear frame by a damper of coefficient  $3.714 \text{ Nm/(rad/s)}$ . This is a practical rotational manifestation of the series damper, inerter, spring network discussed in the references and it resonates at  $52 \text{ rad/s}$ . A translational inerter is illustrated in Figure 15. The results of applying the network to the motorcycle with half the normal frame stiffness are illustrated in Figure 16. The network is just as effective as the dynamic absorber in stabilising the wobble while not de-stabilising the weave, although it is more complex mechanically.



**Figure 15.** Diagrammatic representation of a translational inerter, being a mechanical element giving a force between its terminals which is proportional to the relative acceleration between them (Smith, 2002).

Certain responses to steering torque control inputs are implied by the stability properties discussed. In particular, where the real root passes through zero as speed is increased, as happens with the basic motorcycle at about  $13 \text{ m/s}$ , the zero-order term in the characteristic equation is zero and all the steady-turn responses are infinite (Sharp, 2000; Limebeer and Sharp, 2006). Maximum response to steer torque input ratios were found experimentally by Biral et al. (2003) near to  $8 \text{ m/s}$  speed, apparently corresponding to this stability transition. Also, when an oscillatory mode is lightly damped, it is probable that forcing at the resonant frequency will yield a large response for modest forcing. However, weave and wobble frequencies are high compared with the control bandwidth of a normal person, so rider forcing of wobble or weave is unlikely to play much part in motorcycle steering control.

From any of the linear models, frequency responses to steer torque inputs or to rider lean torque inputs are easy to calculate (Sharp, 2001) but they remain difficult to interpret. Compared with root loci, the response characteristics are relatively complicated. There are several responses potentially of interest, each being a function of frequency



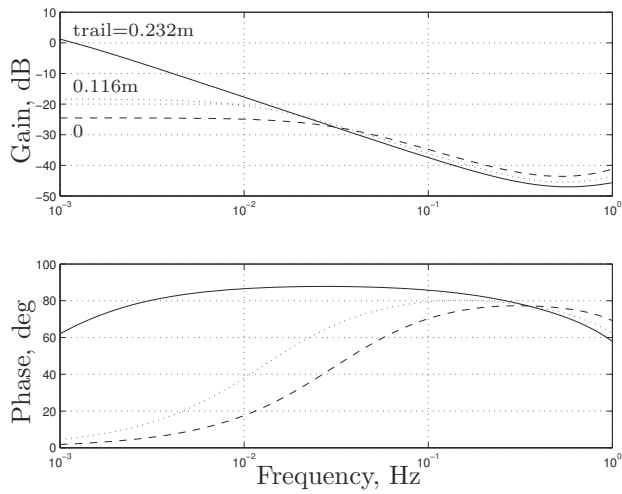
**Figure 16.** Root locus plots for the fully detailed motorcycle of Figure 12, with and without a mechanical network in place of a conventional steering damper, over 1 to 56  $m/s$  speed range. The network is a series spring/mass/damper subsystem with inertia  $0.0624 \text{ kgm}^2$ , stiffness  $169.0 \text{ Nm/rad}$  and damper coefficient  $3.714 \text{ Nm/(rad/s)}$ . Plot symbols indicate speed as in Figure 4. The network is similar to the dynamic absorber in stabilising the wobble mode without worsening the weave mode.

and speed. To avoid information overload, we need to be selective in dealing with them. Firstly, we observe that the responses to rider lean torques are normally relatively small, Aoki (1979); Katayama et al. (1988); Sharp (2001) and we pay no attention to them. Secondly, let us restrict attention to the yaw rate response, as representative of manoeuvring in general. Since motorcycles sideslip very little (Sharp et al., 2004), the yaw rate is closely related to the lateral acceleration, which in turn is closely related to the roll angle for low frequency motions. Thirdly, we consider the rider's control bandwidth to extend to  $1 \text{ Hz}$  only. With these restrictions, yaw rate frequency responses for the fully detailed model of Figure 7 are shown in Figure 17 for  $20 \text{ m/s}$  speed and in Figure 18 for  $40 \text{ m/s}$  speed.

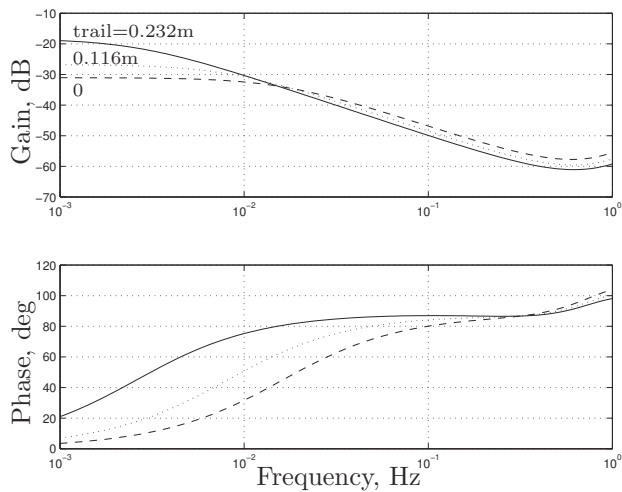
A notable feature of the results is that the rolling-off of the gains with increasing frequency occur at much lower frequencies than would be the case for a car. This feature is well aligned with the known limitations of motorcycles with respect to obstacle avoidance capabilities (Watanabe and Yoshida, 1973). Relationships between design changes, involving trail as shown, wheelbase, wheel spin inertias and steer damping, and these yaw rate responses, for different speeds, are complex. The relationships cannot, at present, be related to handling qualities in any convincing way. However, it can be surmised that motorcycle design modifications which increase the system bandwidth will be advantageous subjectively, other things being equal.

Responses to a unit step input of steering torque for the motorcycle setup of Figure



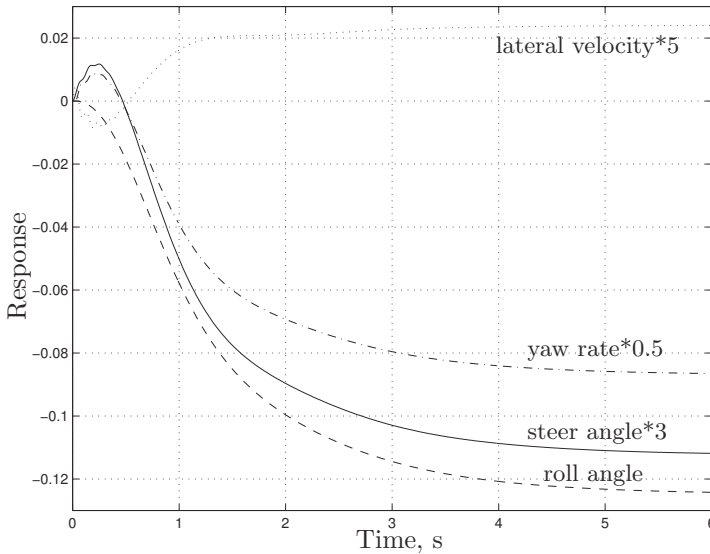


**Figure 17.** Yaw rate frequency responses to steer torque input for the fully detailed motorcycle of Figure 7 at  $20\text{ m/s}$ , with variations in mechanical trail.



**Figure 18.** Yaw rate frequency responses to steer torque input for the fully detailed motorcycle of Figure 7 at  $40\text{ m/s}$ , with variations in mechanical trail.

7 at a speed of  $7\text{ m/s}$ , which is in the auto-stable range, are shown in Figure 19. The motorcycle ends in a steady circular turn to the left, after an initial steer angle response to the right. Despite that the steering torque is maintained throughout, self-steering overcomes the input and the steer angle in the steady turn is to the left. The roll angle response is monotonic, while the lateral velocity response oscillates before settling to a positive value. If the speed chosen were outside the auto-stable range, the rider would have to apply a second torque, counter to the original input, to prevent the machine from diverging in roll, see Limebeer and Sharp (2006).



**Figure 19.** Responses in steer angle, roll angle, yaw rate and lateral velocity of rear tyre to ground contact point following a  $1\text{ Nm}$  step steer torque input with zero initial conditions and a speed of  $7\text{ m/s}$ . The machine is auto-stable at this speed, see Figure 7, so that it settles to a steady turn to the left for a right-hand steer input. Units are SI.

**A summary of the base model and its derivatives.** This sequence of theoretical results supports the following conclusions relating to the representation of the small perturbation motions of a motorcycle. All the models give substantially the same behaviour up to about  $10\text{ m/s}$ , so the basic model, having pure-rolling tyres constraining the available freedoms, is adequate for representing low speeds. Experiments conducted by Schwab et al. (2006) using the speed range  $3$  to  $6\text{ m/s}$  clearly confirm this conclusion for the bicycle case. For higher speeds, the inclusion of sideslip with realistic tyre cornering properties brings increasing departures from the basic motorcycle behaviour as the speed builds. Instead of increasing indefinitely with speed, the weave mode damping reaches a maximum at about  $16\text{ m/s}$  and becomes quite small at high speeds. A wobble mode appears, with damping decreasing towards zero as the speed increases to  $50\text{ m/s}$ .

Including tyre relaxation lags in the model brings decreases in both wobble and weave mode damping factors, where they are already low. At this modelling level, for other than very low speeds, stability problems are expected to occur only at high speeds.

Frame torsional compliance destabilises the weave mode and affects the way the wobble mode varies with speed markedly. Flexible frames give rise to worst-case wobble damping at moderate speeds, which was mysterious when it was observed (Eaton, 1973; Roe and Thorpe, 1976; Dunlop Tyre Company, c1977; Verma et al., 1980) before an explanation was found (Sharp and Alstead, 1980; Spierings, 1981). Spierings (1981) especially showed that gyroscopic torques deriving from the spinning front wheel are influential on wobble and the distribution of the frame compliance in the main frame and front forks is important in the design of a typical machine. A steering damper will add damping to the wobble mode but will subtract damping from the weave mode. There appear to be alternatives to the steering damper which are not fully developed yet. These alternatives include tuned dynamic absorbers and mechanical networks containing inerters, apparently able to damp the wobble without harming the weave or other aspects of the handling behaviour.

Motion responses to steering torque control inputs can be calculated readily. Variations in responses with frequency, speed and design parameter value changes are complicated and are not, at present, convertible into quality judgments that could be used to guide design.

### 2.3 Experimental work

Motorcycle motions have been measured by several researchers (Döhning, 1956; Eaton, 1973; Watanabe and Yoshida, 1973; Jennings, 1974; Roe and Thorpe, 1976; McKibben, 1978; Rice, 1978; Zellner and Weir, 1978; Aoki, 1979; Weir and Zellner, 1979; Verma et al., 1980; Otto, 1980; Thomson and Rathgeber, 1984; Takahashi et al., 1984; Bayer, 1988; Sugizaki and Hasegawa, 1988; Katayama et al., 1988; Biral et al., 2003). A proportion of these studies include comparisons of measured behaviour with model predicted behaviour, using the kind of model described in the earlier sections (Eaton, 1973; Weir and Zellner, 1979; Verma et al., 1980; Thomson and Rathgeber, 1984; Takahashi et al., 1984; Katayama et al., 1988; Biral et al., 2003). Typically the test motorcycle is instrumented with power supplies, sensors, signal conditioners and recording media, and responses to excitation are stored for post-processing (Zellner and Weir, 1978; Biral et al., 2003). Low-frequency excitation can be supplied easily by a test rider, while higher frequency inputs can be obtained by impulsive forcing of the handlebars, dropping and catching again a falling weight or using a lateral-acting rocket motor. A key measurand is the steering torque, which is usually obtained through a strain-gauge device. Accelerometers and rate gyroscopes are especially easy to fit, while rotary (for steer angle) and linear (for suspension deflections) potentiometers are a little more troublesome. Recent advances in solid-state sensing, analog to digital conversion and digital data storage have made the acquisition of good data quite inexpensive. The single most important output to input relationship is probably that between yaw rate and steer torque. Such measurement programmes have established, in general terms, that good theory as outlined above for near-straight-running conditions is closely aligned with good practice.

### 3 High-fidelity modelling

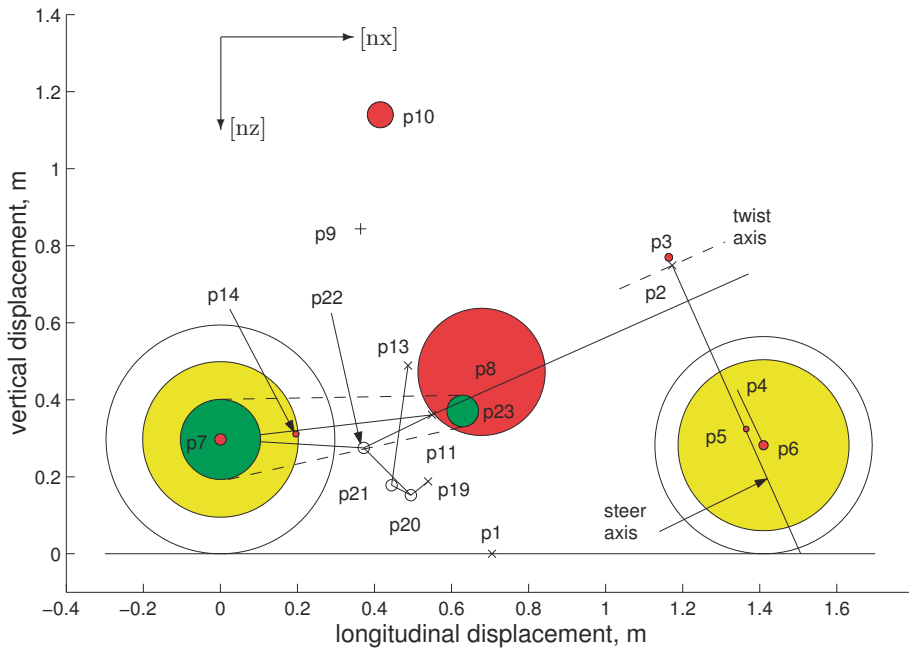
The discussion above concentrates on running at constant speed on a flat road with only small perturbations from straight running. The restrictions on the generality of the motions allowed imply that the longitudinal dynamics and the vertical (or in-plane) dynamics decouple from the lateral (or out-of-plane) dynamics and the three sets of problems can be conveniently separated at this level of treatment. At the simple level also, the suspension system of the motorcycle plays little part in the lateral dynamics and it has so far been omitted from consideration. The tyre forces are amenable to elementary description, since the tyres do not alter their loadings much from the static levels and they operate only in the linear region corresponding to small camber angles and sideslip values. Aligning moments and aerodynamic forces have been neglected also. Models can be built by hand, although it is more convenient, more reliable and faster to use automated multibody software, even for relatively simple models.

In order to deal with more general motions, the modelling problem becomes much larger. All six degrees of freedom for the main frame of the machine need to be allowed, suspension system models including kinematic descriptions of links and transmission elements need to be included. Aerodynamic forces and moments need to be considered and a comprehensive tyre force and moment model is essential. The tyre force model will probably need to allow the ground contact patch to migrate around the tyre circumference and sidewall, accommodating the quite complex geometrical changes that occur at large camber and steer angles. Hand analysis for models at this advanced level are not practical, although the topic is exposed in Pacejka (2002), since they will require huge resource to complete and are unlikely to ever be mathematically correct.

#### 3.1 A contemporary model

A mathematical model that is considered representative of modern high-performance, road-going machines, travelling on flat and level road surfaces, with rider-upper body given a roll degree of freedom relative to the main frame, is illustrated in Figure 20. The model parameters, see (Sharp et al., 2004, 2005a) are based on those of the Suzuki GSX-R1000, which has the common telescopic front fork suspension and a swinging arm rear suspension with monoshock spring/damper arrangement. It is of tree-structure, except for one kinematic-loop closure in the monoshock description. Alternative ways of representing the kinematics of the monoshock suspension are examined in Sharp et al. (2004). Sharp et al. (2005a) also discusses alternative motorcycle architectures and modelling choices that can be exercised, with the resulting efficiency of simulation runs.

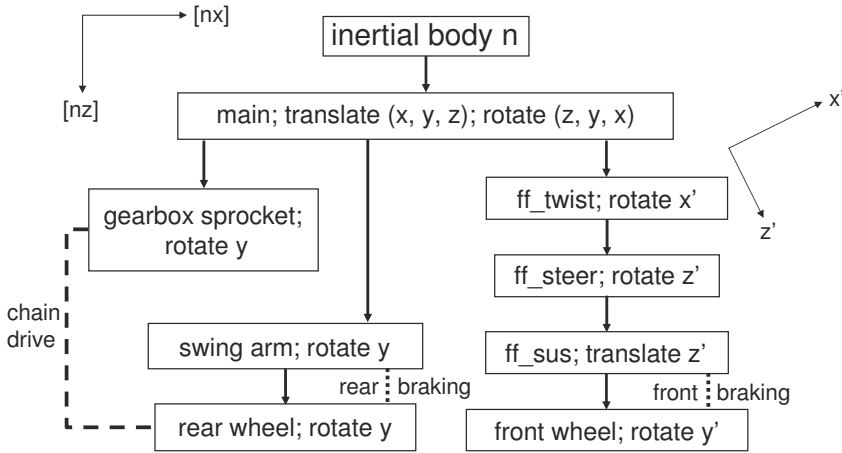
The main frame is allowed unrestricted motion, while the rider's upper body and the rear swing arm are pin-jointed to it. Torsional compliance at the steering head is allowed by the inclusion of a massless frame with roll freedom and spring / damper restraint to the main frame, the twist angle being treated as small. The steering system is pin-jointed to the massless twisting frame at the steering head. The lower front forks and front wheel can translate along the line of the forks, relative to the upper forks, and each road wheel is presumed axi-symmetric and allowed to spin. The bodies included and the freedoms allowed are shown in Figure 21. For the closed loop of links of the monoshock mechanism, the kinematic chain is notionally broken for the definition of the chain. Then



**Figure 20.** Scale representation of a modern high-performance production motorcycle, having telescopic front forks, swing-arm rear suspension with monoshock control and chain drive. Masses are depicted by shaded circles with diameter in proportion to the relevant mass. Points marked by crosses belong to the main frame. p10 corresponds to the mass centre of the rider's upper body and p9 to the horizontal pin-joint for that body. The monoshock unit acts between p13 and p21. p11 is the swing arm pivot.

it is connected together again with the addition of constraints stipulating that chain end points must coincide with each other. Powertrain rotors can be added if desired, with gearing to the spin of the rear wheel. All such rotors add to the effective longitudinal inertia of the motorcycle and affect its acceleration and deceleration but contra-spinning rotors cancel as far as gyroscopic influences on the handling are concerned.

Each tyre has width and the contact points migrate, in general both circumferentially and laterally, so that the lowest point on each one is taken to be the contact centre. The front tyre / road contact geometry is described as follows, see Figure 22. The height of the wheel centre O from the ground is the vertical component of the vector joining it to the origin. The wheel spindle and vertical directions allow the direction of  $OC$  to be found, from which the height of C is known. The contact point G lies directly below C at road level and the distance GC gives the change in the tyre carcass compression from the nominal state and hence the tyre load. The motion state of G is used to calculate the longitudinal and lateral tyre slips and the wheel camber angle comes from the orientation of the wheel spindle vector. The rear tyre is dealt with in the same way but the geometry

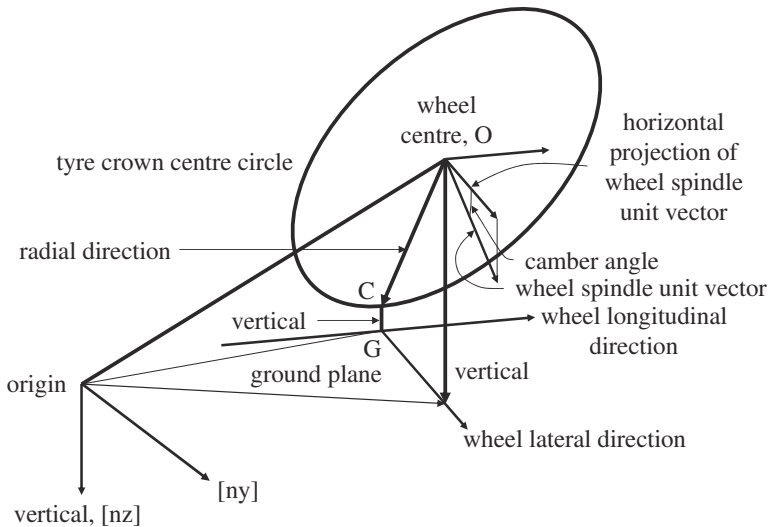


**Figure 21.** Multibody structure and freedoms allowed in contemporary high-fidelity motorcycle model (Sharp et al., 2004, 2005a).

is much simpler in that case.

The tyre load, longitudinal slip, lateral slip and camber angle provide the necessary inputs to the Motorcycle Magic Formula (Pacejka, 2002; Sharp et al., 2004), which requires an extensive data set to define a particular tyre. Negative tyre loads are disallowed and all tyre forces and moments go to zero as the tyre load does, so that wheelies and stoppies can be simulated properly. Forces and moments due to aerodynamic influences, suspension elements, frame flexing and rider movements are added. Suspension springs and dampers are defined as linear but suspension and steering limit stops are included. The transmission chain is treated as an elastic damped tension link joining the sprockets together, see Figure 20. Either the upper chain run or the lower one may be in tension but not both together. A small clearance may be specified to represent slack in the chain.

Since uncontrolled motorcycles are unstable under some running conditions, it may be necessary in some simulation contexts to incorporate stabilising feedback steering control. With such control, together with a speed controller, roll angle and speed targets can be tracked, enabling the finding of equilibrium cornering (trim) states. Speed-adaptive proportional, integral, derivative action steering control and proportional, integral action drive control schemes are described in Sharp et al. (2004). The influences of acceleration or braking can be represented nicely, while preserving the convenience of a constant-speed model, by adding an artificial horizontal gravitational field (Hales, 1965; Sharp, 1976; Limebeer et al., 2001a,b). This device reproduces load transfers and the inertial effects of changing speed on each body mass centre, and allows the computation of eigenvalues that align closely with the corresponding ideas that come from simulation runs (Limebeer et al., 2001b). Road profiling of long wavelength in comparison with the



**Figure 22.** Front wheel geometry, illustrating how the contact centre  $G$  is found, and how its position and velocity feed into the calculation of tyre forces and moments.

tyre contact length can be included, by specifying a varying elevation for the ground at front and rear contact points, influencing the calculation of tyre loads and shear forces. Steering control inputs can be considered to consist only of handlebar torque or to also include rider-upper-body-lean torque. The model is considered valid, in a generic sense, for the full range of feasible speeds, wheel loads and lean angles, tyre slip angles less than  $8^\circ$ , and vibration frequencies less than about 18 Hz.

### 3.2 Model exploitation

The multibody software used for model building, AutoSim, allows the model file to be loaded in either of two ways, linear or nonlinear. In the former (linear) case, the system is symbolically linearized, for small perturbations about a general trim condition, and a MATLAB “M” file can be written automatically, to set up the linear model in state space form. The symbolic MATLAB file can be used for the complete range of linear system operations encompassed by the software. Such operations include automated conversion of the linear model to discrete-time form and discrete linear quadratic regulator optimization. The linear model generated by AutoSim has 27 states, including the longitudinal displacement from the origin and the wheel spin angles. In the latter (nonlinear) case, a ready-to-compile-and-link simulation program can be obtained. The model would need additional features to enable it to simulate travel over a generally profiled three dimensional surface, to follow a prescribed path along a roadway or to respond to a cross-wind disturbance. A generally profiled road is included in a commercial development of the model family described, BikeSim (see <http://www.carsim.com>), and the path-following

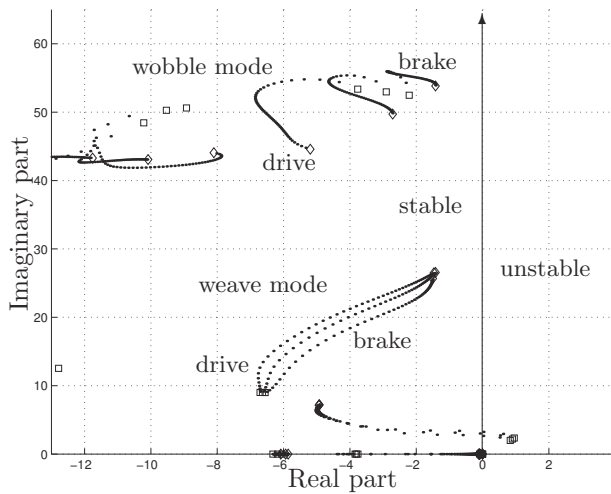
problem is discussed in section 5 on page 28.

## 4 Behaviour predicted by advanced models

The kind of high-fidelity model described above has a very wide range of capabilities to simulate and to otherwise analyse motorcycle motions. Standard computations include the following: (i) trim states corresponding to full ranges of speed, roll angle and acceleration (Limebeer et al., 2002; Pacejka, 2002; Sharp et al., 2004); (ii) root loci for small perturbations from any of the trim states of interest (Sharp, 2001; Pacejka, 2002; Limebeer et al., 2002; Sharp et al., 2004); (iii) mode shapes belonging to any of the roots calculated (Sharp, 2001; Pacejka, 2002; Sharp et al., 2004); (iv) eigenvalue sensitivities to design parameter variations (Sharp, 1994, 2001); (v) frequency responses to steering torque and body lean torque control inputs (Sharp, 2001; Cossalter, 2002); (vi) steer angle responses to road undulation forcing for the cornering machine (Limebeer et al., 2002; Sharp et al., 2004); (vii) nonlinear responses to large amplitude regular road forcing (Shaeri et al., 2004; Sharp et al., 2005b); (viii) nonlinear stability analysis and bifurcations of solutions (Meijaard and Popov, 2004, 2006); and (ix) transient motions following changes in control settings (Pacejka, 2002; Cossalter, 2002; Sharp et al., 2005a).

To illustrate just one of these categories, let us examine the influences of acceleration or deceleration on the stability of the motorcycle of Sharp et al. (2004) when it is cornering with a sustained  $20^\circ$  roll angle through a range of speeds. Firstly, the rider-stabilised machine is employed to find the relevant trim states. To deal with accelerating or decelerating, an artificial horizontal gravitational field is introduced, aligned with the longitudinal axis of the motorcycle. Then an extremely small acceleration is specified over a long simulation run, so that the speed range of interest is covered in the run. In the simulation run, it is as if “inertia forces” were acting on each body’s mass centre in opposition to the specified acceleration, as discussed in subsection 3.1 on page 23. Each simulation run establishes the trim states for a given level of acceleration, here set at  $-2.5\text{ m/s}^2$  to represent engine driving, zero for constant speed, and  $2.5\text{ m/s}^2$  to represent braking by the rear wheel and aerodynamic drag. The appropriate trim state is used to create the linearised model for small perturbations from that state and the eigenvalues of the linearised model found. The resulting root loci are shown in Figure 23, where the speed range for braking and constant speed is 4 to  $67\text{ m/s}$ , while that for driving is 4 to  $57\text{ m/s}$ , curtailed due to the difficulty of stabilising the machine at high speeds with the rather rudimentary feedback steering controller used. Accelerating and braking do little to affect the stability of the weave mode at high-speed but braking destabilises the wobble mode, while accelerating stabilises it. Apparently, opening the throttle to damp down a cornering wobble oscillation may be effective if the vibrations are detected early enough.





**Figure 23.** Root locus plots for the motorcycle of Figure 20, described in Sharp et al. (2004), when braking, at constant speed, or accelerating at a steady roll angle of  $20^\circ$ . The speed range for the first two cases is 4 to  $67\text{ m/s}$ , while for the third, it is 4 to  $57\text{ m/s}$ . The influences of acceleration on the high-speed weave are small but braking destabilises the wobble mode, while accelerating stabilises it for these running conditions. The symbol  $\square$  marks the low-speed end, while the  $\diamond$  marks the high-speed end of each locus.

## 5 Rider control and handling quality issues

### 5.1 General observations

In principle a rider may operate in “fixed control”, effectively using steer angle and/or upper body lean angle as controls or alternatively, he/she may use “free control” in which steer torque and lean torque are the control inputs. The same options exist in relation to longitudinal control by brake or throttle actions but conventional arrangements more or less compel the rider to use brake force control, since the levers move only a little, and it is difficult to distinguish between throttle position control and throttle torque control since the position and the torque are very strongly correlated (Sharp, 2000). Lateral control is more complex and interesting than longitudinal control, so the discussion will focus on steering and on free control, since that is bound to be the rider’s preference, for reasons given in section 1 on page 1.

It is evident that riders are learning and optimising controllers. They learn by repeated trials and observations and they optimise some combination of system performance like path following accuracy, speed maintenance and control effort. In a learning phase, control strategies are evaluated. The control mode and the concentration level chosen will vary with the circumstances from high effort, high accuracy at the one extreme to the converse. In the trials, the relative effort and reward associated with body torque and steer torque controls will be assessed, it being apparent, at the outset, that

using body muscles for active control is likely to be feasible only at very low frequencies. For most people, a realistic estimate of the bandwidth for body control is perhaps 0.1 Hz, rather than the 1 Hz associated with steer torque control (Aoki, 1979; Sugizaki and Hasegawa, 1988). It is surmised, at this stage, that riders use feedback control for stabilisation and disturbance rejection, and preview control for path following and speed maintenance. Provided that the available steering torque and/or rider-body-lean torque actuations are effective, as will normally be the case, stabilisation of divergent motions is possible if the time constant is sufficiently long, and stabilisation of oscillatory modes by active control is possible if the frequency of the oscillations is low enough in relation to the control bandwidth.

The rider can also use near zero frequency control to effectively alter the higher frequency dynamics of the motorcycle-rider system. When the rider substantially alters his body position relative to the motorcycle for drag reduction or vigorous cornering, it is best viewed as parameter adaptation, not as active control. Such adjustments are most obvious in racing and they typically occur well before a manoeuvre. The equilibrium or trim states of the system are affected and the dynamics involving small perturbations from trim are correspondingly influenced. When the rider finds that the changes are favourable, reinforcement learning (Gurney, 1997) will contribute to making them into habits.

Sometimes, it is argued that high stability margins lead to poor manoeuvrability (Cossalter, 2002) and therefore that stability can be excessive. However, manoeuvrability depends fundamentally on the relationship between the external forces that can be called upon by the rider and the responsiveness of the motorcycle to those forces. This is substantially independent of the stability properties. The rider has more or less direct control of the lateral force at the front tyre contact with the ground. From a straight and level condition, maximum response is obtainable by steering the front wheel quickly to the condition at which it will give maximum sideforce. Most riders are not brave enough or skillful enough to do this. Typically, they will extract much less than the full performance capability of their machine through conservative riding (Rice, 1978). Then, the extent to which the machine will respond to the external forces imposed will depend on its mass and inertia properties, its wheelbase and many design details, combining in a complex way that can be understood only by examining the full nonlinear system dynamics at limiting performance. It seems likely that good stability margins will actually make a positive contribution to manoeuvrability, by encouraging riders to fully use the performance envelope, without fear of losing control.

## 5.2 Rider modelling

It can be anticipated that successful modelling of rider control activity will yield some understanding of how the rider works and of how to improve rider / machine interactions. Rider models are also clearly useful in allowing path-based simulation runs to be made. A simple approach to rider modelling is to first provide the kind of stabilising feedback control described in section 3.1 on page 23 and to add a single-point preview control. Such a control is based on the idea that the rider uses a notional point in the motorcycle, but well ahead of it, as a reference point, with which to compare the intended path at the same

longitudinal location. The lateral distance between the notional vehicle point and the path point is used as an error signal, and the error is employed in a proportional / integral / derivative control scheme (say) to provide steer torque. Such a scheme would have four parameters, the preview distance and three gains, which could be chosen by heuristic methods or by a parameter optimisation method to give a reasonable performance. The obvious difficulty, even if the parameters are optimised, is that the path information used by the rider is artificially restricted, to the point where it is likely to strongly prejudice the control quality. If the path slope and curvature at the preview point are also used for control, the information available to the controller is improved but the parameter setting problem is becoming significantly more complex, with nine gains instead of the original three. The path information would remain quite restricted compared with what we can expect a real rider to utilise. Apparently, it is easier to achieve a path-based simulation capability than it is to model a rider with sufficient realism to gain some understanding of how that rider operates.

However, suppose we have an optimal control scheme in which full path information is available to the rider for control, and the objective of the optimisation is to minimise path tracking errors, consistent with using acceptable levels of control effort. It seems likely that the control will be human-like, since the rider himself is an optimising system with the same objectives. Within the mathematically linear world, such a scheme exists (Sharp and Valtetsiotis, 2001; Sharp, 2005, 2006), and it will now be described.

### 5.3 Optimal linear preview control

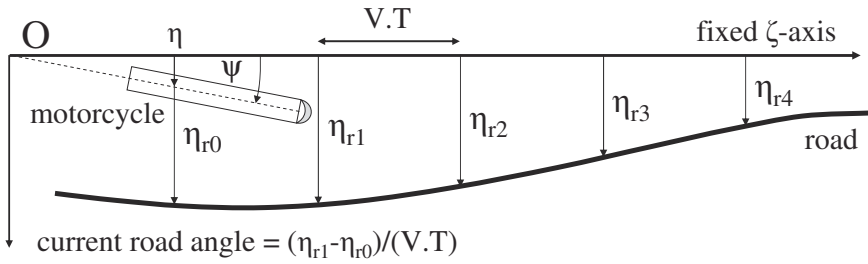
The linear motorcycle / rider model is arranged to use earth-fixed reference axes and to have steer torque and rider-upper-body-lean torque as inputs and lateral displacement of the origin point as an output. The model is then transformed into discrete-time form, typically using a time step of  $0.02\text{ s}$ . Through each time step, the motorcycle travels  $0.02V\text{ m}$  where  $V$  is the speed. A roadway lateral profile is defined by discrete points  $1\text{ m}$  apart longitudinally (say) in the same earth-reference system, so that all the road profile points in front of the motorcycle approach the machine by  $0.02V\text{ m}$  through each time step. In this earth-fixed reference system, illustrated in Figure 24, the road dynamics are those of a shift register and the equations describing these dynamics are of the same form as the equations of the motorcycle. The two sets of equations can be combined to yield a composite system, which, at this stage, has no coupling between the parts.

A cost function is then devised, to express the objective of the system as the achievement of a balance between low path-tracking errors and low expenditure of control power, the balance to be specified by weighting constants. The cost function couples the motorcycle and roadway dynamics, and the equations define a stabilisable, detectable linear discrete-time system amenable to optimisation by standard linear quadratic regulator (LQR) methods.

Suppose that the vehicle equations are:

$$x_v(k+1) = \mathbf{A}_v x_v(k) + \mathbf{B}_v T(k) \quad (5.1)$$

$$y_v(k) = \mathbf{C}_v x_v(k) + \mathbf{D}_v T(k) \quad (5.2)$$



**Figure 24.** Diagrammatic representation of the motorcycle tracking a roadway path, with the whole system referenced to ground. Such a description implies that the road sample values pass through a shift register operation at each time step. The dynamics of the shift register are easy to specify mathematically.

in which  $x_v$  is the vehicle state vector and  $T$  is the torque input, and the road equation is:

$$\eta_r(k + 1) = \mathbf{A}_r \eta_r(k) + \mathbf{B}_r \eta_{ri}(k) \tag{5.3}$$

where  $\eta_r$  is the road state and  $\eta_{ri}$  is the road sample value that enters the system at time  $kT$ . To represent the road shift register process,  $\mathbf{A}_r$  has the form:

$$\mathbf{A}_r = \begin{bmatrix} 0 & 1 & 0 & \dots & \dots & 0 \\ 0 & 0 & 1 & \dots & \dots & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & 0 \end{bmatrix} \tag{5.4}$$

and  $\mathbf{B}_r$  has the form:

$$\mathbf{B}_r = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \dots \\ 1 \end{bmatrix} \tag{5.5}$$

Combining vehicle and road equations together, the full dynamic system is defined by:

$$\begin{bmatrix} x_v(k + 1) \\ \eta_r(k + 1) \end{bmatrix} = \begin{bmatrix} \mathbf{A}_v & 0 \\ 0 & \mathbf{A}_r \end{bmatrix} \begin{bmatrix} x_v(k) \\ \eta_r(k) \end{bmatrix} + \begin{bmatrix} \mathbf{B}_v \\ 0 \end{bmatrix} T(k) + \begin{bmatrix} 0 \\ \mathbf{B}_r \end{bmatrix} \eta_{ri}(k) \tag{5.6}$$

which takes the standard form:

$$z(k + 1) = \mathbf{A}z(k) + \mathbf{B}u(k) + \mathbf{E}\eta_{ri}(k) \tag{5.7}$$

$$y(k) = \mathbf{C}z(k) + \mathbf{D}u(k) \quad (5.8)$$

If  $\eta_{ri}$  is a sample from a white-noise random sequence, the time-invariant optimal control which minimises a quadratic cost function  $J$ , given that the pair  $(\mathbf{A}, \mathbf{B})$  is stabilisable and that the pair  $(\mathbf{A}, \mathbf{C})$  is detectable (Anderson and Moore, 1971) is:

$$u^*(k) = -Kz(k) \text{ where } K = (\mathbf{R} + \mathbf{B}^T\mathbf{P}\mathbf{B})^{-1}\mathbf{B}^T\mathbf{P}\mathbf{A} \quad (5.9)$$

given that the cost function  $J$  is:

$$J = \lim_{n \rightarrow \infty} \sum_{k=0}^n \{z^T(k)\mathbf{Q}z(k) + \mathbf{T}^T(k)\mathbf{R}\mathbf{T}(k)\} \quad (5.10)$$

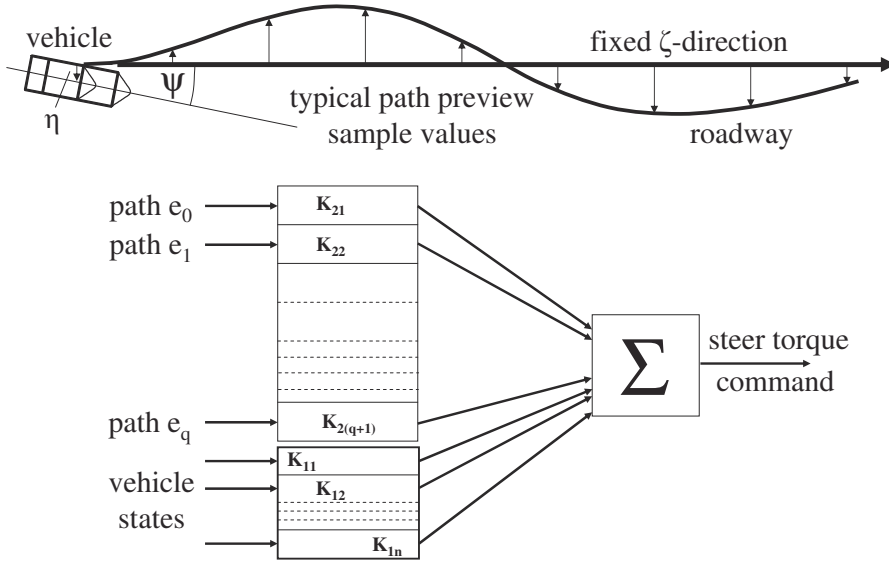
and  $\mathbf{P}$  satisfies the matrix-difference-Riccati equation:

$$\mathbf{P} = \mathbf{A}^T\mathbf{P}\mathbf{A} - \mathbf{A}^T\mathbf{P}\mathbf{B}(\mathbf{R} + \mathbf{B}^T\mathbf{P}\mathbf{B})^{-1}\mathbf{B}^T\mathbf{P}\mathbf{A} + \mathbf{Q} \quad (5.11)$$

where  $\mathbf{Q} = \mathbf{C}^T\mathbf{q}\mathbf{C}$  with  $\mathbf{q}$  a diagonal weighting matrix,  $\text{diag}[q_1, q_2, \dots]$ , with terms corresponding to the number of performance aspects contributing to the cost function, and  $\mathbf{R} = \text{diag}[r_1, r_2]$  has a corresponding relationship to the two control inputs, steer torque and rider lean torque.  $q_1$  is set to represent the importance of path tracking errors,  $q_2$  relates to the absolute lean angle of the rider's upper body, while  $q_3$  refers to the rider's upper body angle relative to the motorcycle frame.  $q_2$  and  $q_3$  values can be chosen to maintain a desirable upper body attitude.  $r_1$  and  $r_2$  are set to reflect the costs associated with steering torque and rider lean torque respectively. It is also known that the optimal control law given by equation 5.9 remains optimal if the white-noise sequence representing the road lateral profile is low-pass filtered, to make it more representative of real roads, when "full" preview of the road (Sharp and Valtetsiotis, 2001; Sharp, 2005, 2006) is available. As explained in Sharp and Valtetsiotis (2001) and more fully in Sharp (2007), the optimal control is best found, using MATLAB, by defining all the matrices in equations 5.7 to 5.10 numerically, partitioning the problem into non-preview and preview parts, yielding:

$$\mathbf{P} = \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12} \\ \mathbf{P}_{21} & \mathbf{P}_{22} \end{bmatrix}$$

and finding the non-preview standard LQR solution first, giving  $\mathbf{P}_{11}$ . As in Louam et al. (1988); Prokop and Sharp (1995); Sharp and Valtetsiotis (2001); Sharp (2007), successive columns of  $\mathbf{P}_{12}$  are determined by adding a new column of the appropriate partition of the cost matrix to the previous column of  $\mathbf{P}_{12}$  premultiplied by the state-transition-matrix of the motorcycle state-feedback closed-loop system. The preview gain vector follows straightforwardly from  $\mathbf{P}_{11}$  and  $\mathbf{P}_{12}$ . The gains fall to zero as the preview distance increases, so that the recursion can be stopped when the new gains found are too small to be interesting. Such a stopping criterion effectively defines "full" preview. More preview than this brings no advantage. Path information far ahead of the vehicle will only be useful at a later time. The structure of the time-invariant optimal control, which represents optimisation over infinite time, is shown in Figure 25.



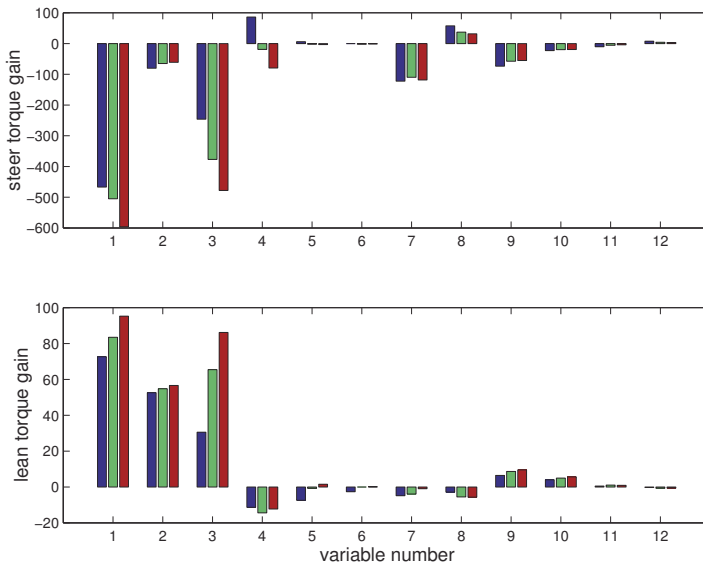
**Figure 25.** Illustration of the time-invariant optimal controls, corresponding to an infinite optimisation horizon and a white noise disturbance. The control involves state feedback with the system state including both vehicle and roadway contributions.

### 5.4 Optimal control results

To develop a set of optimal controls, the nonlinear, feedback-stabilised motorcycle / rider model is run in simulation mode to establish the trim state of interest. That trim state is then used in the linearisation of the system for small perturbations from trim. The linear system resulting is converted to discrete-time form and joined to the shift-register road model as indicated above. Although the process is general, in that the linearisation of the system can relate to any trim state, we will restrict attention here to straight-running trim states, for which the assumed left-right symmetry of the machine and rider brings some reduction in complexity. In particular, the symmetry implies that any feedback gain in the optimal control that relates to an in-plane state variable, like the front-fork extension or the swing-arm angle, will be zero. This feature corresponds to the decoupling of in-plane and out-of-plane motions for small perturbations from straight-running referred to in section 3 on page 23.

The motorcycle will lift and pitch, the tyre loads will change, suspensions will deflect and the rear tyre thrust will work against the aerodynamic drag, and therefore change its small-perturbation properties, as the motorcycle speed changes. These effects are accounted for in the linearisation. Suppose we choose speeds of 10, 35 and 60 *m/s* as representative. We also need to choose the relative weights associated with tracking errors and rider upper-body inclination and with steer torque and lean torque provision, relating to control power. Let us use  $q_1 = 10^4$  and, somewhat arbitrarily at this stage,  $q_2 = 4 * 10^4$ ,  $q_3 = 0$ , to represent tight control by the rider and  $q_1 = 100$ ,  $q_2 = 400$  and  $q_3 = 0$  to represent loose control, and set  $r_1$  and  $r_2$  to be the same, to reflect penalising steer torque and rider-body torque equally, as in Sharp (2006). For  $q_1 = 10^4$ ,  $q_2 = 4 * 10^4$ ,  $q_3 = 0$ ,  $r_1 = r_2 = 1$  and the three speeds 10, 35 and 60 *m/s*, feedback gains are shown

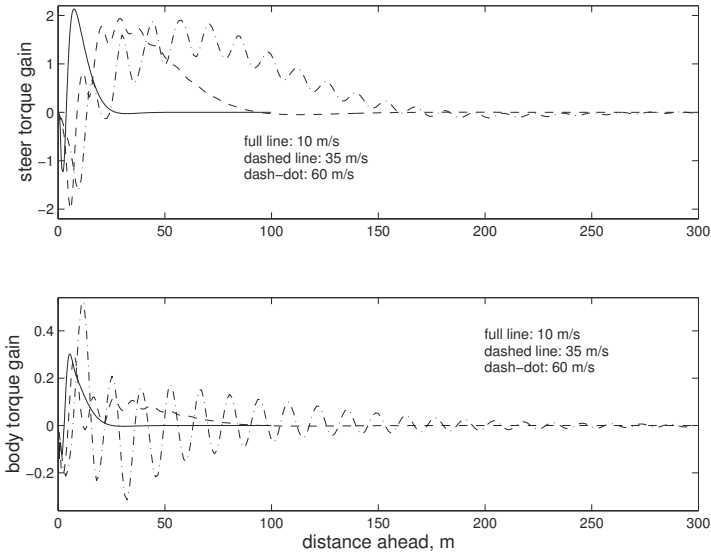
in Figure 26 and preview gains are shown in Figure 27. The state variables used for feedback are (1) roll angle; (2) rider relative roll angle; (3) frame-twist angle; (4) steer angle; (5) rear-tyre sideslip; (6) front-tyre sideslip; (7) lateral velocity; (8) yaw rate; (9) roll velocity; (10) rider-relative-roll velocity; (11) frame-twist velocity and (12) steer velocity. A positive gain implies that a positive value of the variable involved will yield a negative torque, see equation 5.9.



**Figure 26.** Optimal feedback gains for detailed motorcycle model at 10, 35 and 60  $m/s$  with  $q_1 = 10^4$ ,  $q_2 = 4 * 10^4$ ,  $q_3 = 0$  and  $r_1 = r_2 = 1$ . In each group of three bars, the left-hand one is for 10, the centre one is for 35 and the right-hand one is for 60  $m/s$ . Variables are specified in subsection 5.4 on page 33.

In the feedback control, there is a marked tendency for the gain associated with steering torque to be opposed to that for rider-lean torque. For the assumed equal weights on steer torque and body-lean torque, the gains related to the former are an order of magnitude larger than those for the latter. The largest gains relate to roll angle and frame-twist angle and sometimes a gain changes sign as the speed changes from 10 to 35  $m/s$ . This is true of the steer-torque gain number 4, which is equivalent to a steering spring at low speeds and a negative spring at higher ones. The steer-torque gain number 12 is like a steering damper of conventional sign, but rather small. Steer-torque gains numbers 1 and 9 are like the proportional and differential components respectively in the empirical stabilising control mentioned in subsection 3.1 on page 23.

Preview control gains are shown as continuous curves but, for each case, they are actually a succession of discrete values with spacing  $0.02V m$ . Each gain sequence converges to zero as the preview extends away from the motorcycle, confirming the well-known feature of preview control that far future information is not useable to advantage (Tomizuka,



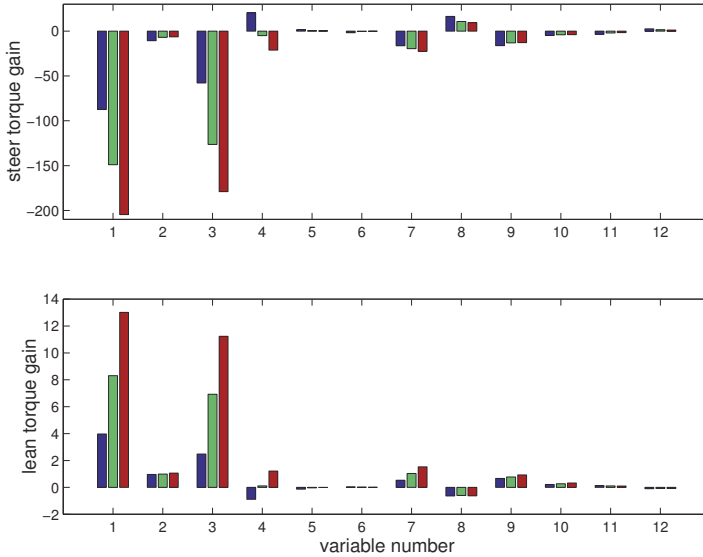
**Figure 27.** Optimal preview gains for detailed motorcycle model at 10, 35 and 60  $m/s$  with  $q_1 = 10^4$ ,  $q_2 = 4 * 10^4$ ,  $q_3 = 0$  and  $r_1 = r_2 = 1$ . Gains are discrete values at  $0.02V$   $m$  intervals.

1976). The 400 preview points assumed are more than sufficient for the full benefits of preview to be realised. With the tight control implied by  $q_1 = 10^4$ ,  $q_2 = 4 * 10^4$ ,  $q_3 = 0$  at 10  $m/s$  the rider needs to look ahead by about 30  $m$ , at 35  $m/s$  the corresponding need is for 100  $m$  preview, and at 60  $m/s$  the rider requirement is about 300  $m$  preview. The “full” preview times are 3, 3 and 5  $s$  respectively, the “full” preview time increasing with speed as found in other cases, being much larger for a motorcycle than for a car at the same speed (Sharp and Valtetsiotis, 2001; Sharp, 2005, 2006). The most useful path information is 7, 30 and 60  $m$  ahead for the speeds 10, 35 and 60  $m/s$  respectively, close to 1  $s$  in time in each case.

Corresponding results for loose control, as reflected in the values  $q_1 = 100$ ,  $q_2 = 400$ ,  $q_3 = 0$ , are shown in Figure 28 relating to state feedback and in Figure 29 relating to preview. Feedback gains show the same pattern as for tight control but they are about one third as large for steer torque and only one tenth as large for rider body torque. Preview gains are approximately 100 times smaller, the most used information is further from the motorcycle at 12, 70 and 130  $m$  away for speeds 10, 35 and 60  $m/s$  respectively, and “full” preview demands significantly more preview time than before. Optimal controls generated for  $r_2 = 0.1$ , corresponding to a lower penalty on the use of body-lean-torque control, have shown that larger gains for body torque, roughly in proportion to the weight reduction, result. Only very small changes occur in the steer torque laws. The theory is in agreement with the theoretical conclusions of Weir and Zellner (1978) and the experimental findings of Aoki (1979); Katayama et al. (1988)



that body lean control does not contribute much to overall directional control of the motorcycle.

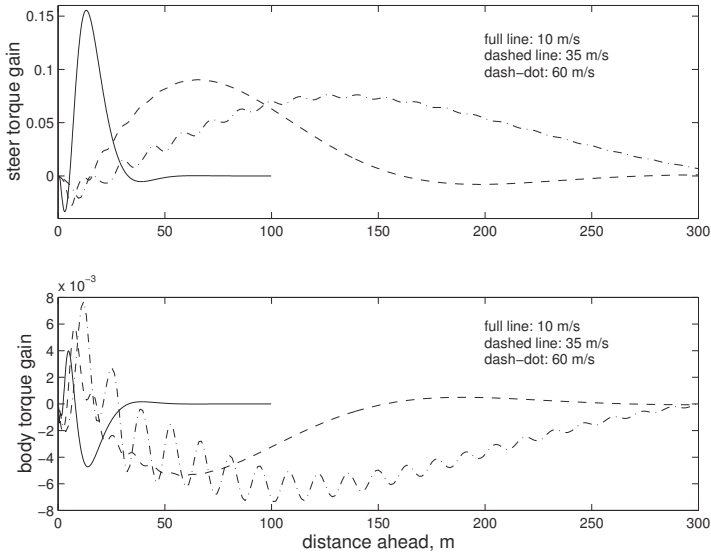


**Figure 28.** Optimal feedback gains for detailed motorcycle model at 10, 35 and 60  $m/s$  with  $q_1 = 100$ ,  $q_2 = 400$ ,  $q_3 = 0$  and  $r_1 = r_2 = 1$ . In each group of three bars, the left-hand one is for 10, the centre one is for 35 and the right-hand one is for 60  $m/s$ . Variables are specified in subsection 5.4 on page 33.

The preview distance needed for full performance benefit in a particular set of running conditions can be quantified rather crudely by finding the point along the relevant preview gain sequence for which the area under the gain curve amounts to 99.9% of the total area out to infinity. Such preview distances are shown as functions of travel speed in Figure 30 for tight controls in the upper part and for loose controls in the lower part. In each case, results for  $r_1 = r_2 = 1$  and for  $r_1 = 1$  and  $r_2 = 0.1$  are included. The necessary preview increases more than in proportion to speed, with all cases being similar. For tight controls, the requirements for body-lean torque and for steer torque are not exactly the same at high speeds and it is not known why this is so, but it does not seem important, since steer torque control is completely dominant. In any case, the pattern may change with variations in the relative weightings of  $q_1$ ,  $q_2$  and  $q_3$ . Again, these results highlight the contrast between motorcycles and cars in respect of road preview needed as a function of speed (Sharp, 2005, 2006).

### 5.5 Path tracking results using the optimal controls

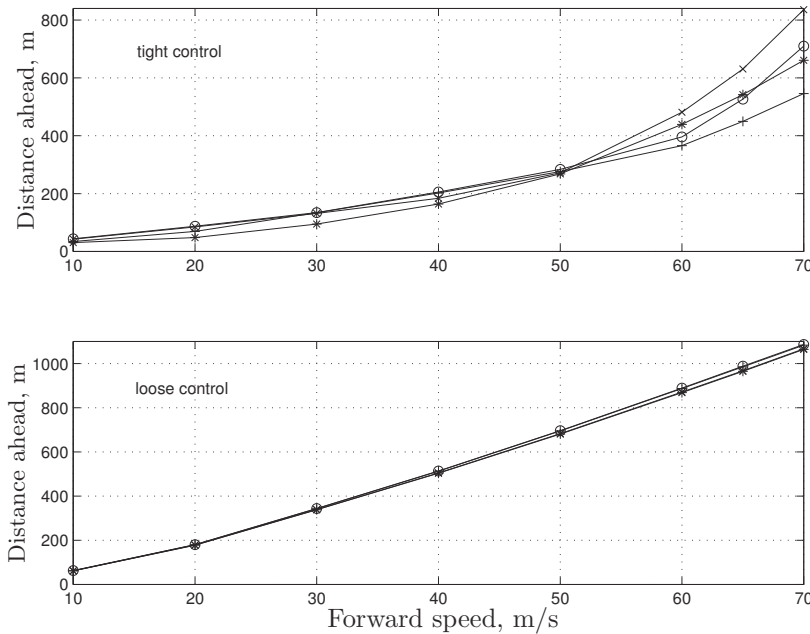
The simplest approach to path tracking is to use the optimal controls developed in the fixed reference frame described above, to close the loop around the motorcycle / road system and to update the state through time, by repeatedly multiplying it by the closed-



**Figure 29.** Optimal preview gains for detailed motorcycle model at 10, 35 and 60  $m/s$  with  $q_1 = 100$ ,  $q_2 = 400$ ,  $q_3 = 0$  and  $r_1 = r_2 = 1$ . Gains are discrete values at  $0.02V m$  intervals.

loop state-transition-matrix. At each update, a new road sample, from the furthest point along the preview distance, enters the problem as input. This simple approach implies that the restriction to small angles in the linear theory used to find the optimal control carries over into the tracking simulation. Paths which diverge substantially from the starting direction will be followed with imprecision and not followed at all if the path direction changes by  $90^\circ$  or more. A better approach is to transform the problem to the local frame of the vehicle, that is, to take the rider's view rather than a ground observer's view of the problem. With such a transformation, at each time step, the reference axes are re-located to align with the motorcycle and the roadway is described relative to the motorcycle. Lateral and yaw position feedback terms are lost from the control scheme, because there are no corresponding errors, due to the re-positioning of the reference axes. The transformation of the path data exactly compensates for the loss of these terms, as shown in Sharp and Valtetsiotis (2001). The small-angle theory now requires only that the path direction does not change radically over the preview distance. Bearing in mind that sharp turns are practical only for low speeds and that for low speeds, relatively short previews are necessary, the system adapts to the circumstances such that small-angle theory is invariably satisfied when a realistic manoeuvre is being made.

The defined path starts with a straight section and the road-preview states of the full system are loaded initially with zeros. This beginning state corresponds with the motorcycle having travelled along a straight path, before the simulation commences, for the duration of the preview. The path lateral profile is then “seen” in the distance first,

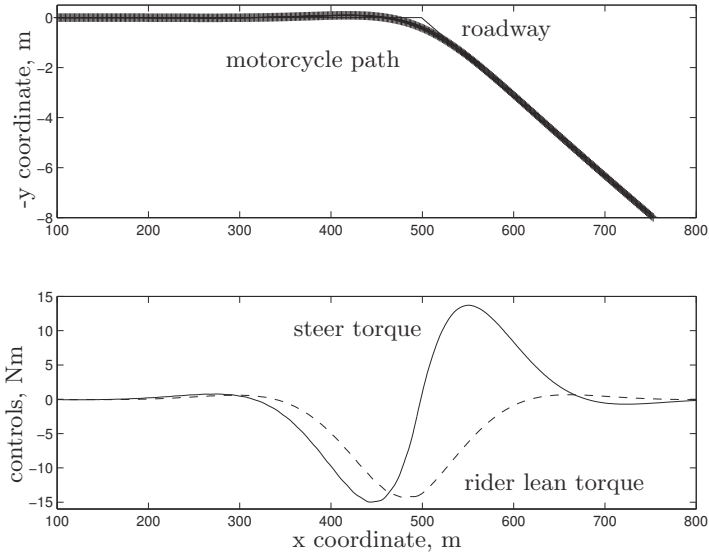


**Figure 30.** Preview distances required for full performance according to 99.9% area criterion, as functions of travel speed. The upper graph is for tight controls,  $q_1 = 10^4$ ,  $q_2 = 4 * 10^4$ ,  $q_3 = 0$ , while the lower one is for loose controls,  $q_1 = 100$ ,  $q_2 = 400$ ,  $q_3 = 0$ . The plot symbols  $\circ$  and  $+$  relate to steer torque, while  $\times$  and  $*$  relate to body-lean torque.  $\circ$  and  $\times$  refer to cases  $r_1 = r_2 = 1$  and  $+$  and  $*$  refer to cases  $r_1 = 1$  and  $r_2 = 0.1$ .

as with a rider with a full view of the road ahead. The simulation proceeds by stepping through time, using the step length  $0.02\text{ s}$  employed for the discrete-time optimal control calculations, until a specified number of steps has been completed. At each step, the absolute position and orientation of the motorcycle is updated and the ground-based road lateral profile data in front of the machine is read in, transformed to the local view of the rider and used to replace the old values in that part of the system state vector that corresponds to the road profile. Spline interpolation is used to obtain the local lateral road profile at the proper preview points. At the same time, the motorcycle states representing its lateral position and its attitude angle are set to zero, corresponding to the reference axes being re-located to align with the machine, (Sharp and Valtetsiotis, 2001; Sharp, 2005, 2006).

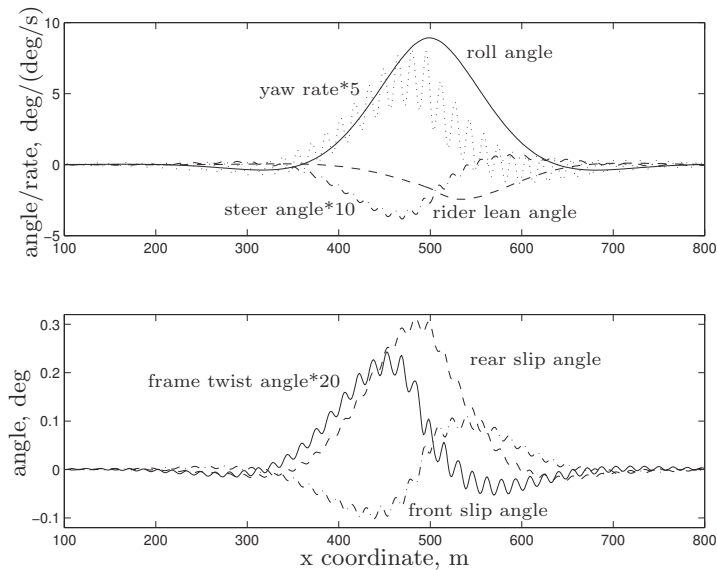
Many manoeuvres have been studied, including randomly generated, low-pass-filtered lateral profiles as in Sharp and Valtetsiotis (2001), lane changes with geometry from Rice (1978), sudden direction changes as in Sharp and Valtetsiotis (2001); Sharp (2005) and an S-path manoeuvre devised by Frezza and Beghi (2003), see (Sharp, 2006). Some path-tracking simulation results are illustrated next. Suppose we choose a sudden  $15^\circ$  direction change manoeuvre at the high speed of  $70\text{ m/s}$  with  $500\text{ m}$  preview, tight control, and

equal weightings on steer and body lean torques, represented by  $q_1 = 10^4$ ,  $q_2 = 4 * 10^4$ ,  $q_3 = 0$  and  $r_1 = r_2 = 1$ . The high speed and the tight control cause the preview control gain sequence to be oscillatory, as apparent in Figure 27. Tracking quality and control inputs are shown in Figure 31. A selection of corresponding motion variables is shown in Figure 32 and the tyre lateral forces are shown in Figure 33.



**Figure 31.** Tracking simulation results for a sudden  $15^\circ$  direction change and  $70\text{ m/s}$  speed with  $500\text{ m}$  preview, tight control and equal weights on steer torque and rider lean torque ( $q_1 = 10^4$ ,  $q_2 = 4 * 10^4$ ,  $q_3 = 0$ ,  $r_1 = r_2 = 1$ ). The upper part shows the path of the motorcycle and the lower part shows the control torques.

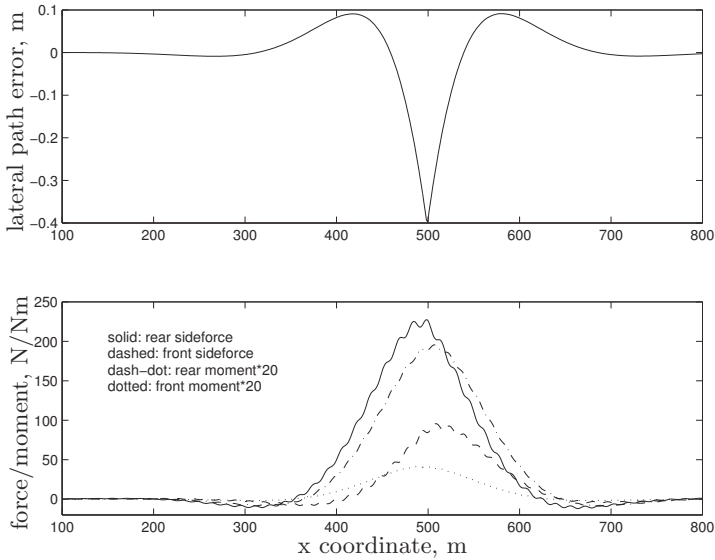
Despite the tight control employed, some corner-cutting occurs. The rider's body lean torque has a peak value similar to that of the steer torque, occurring at the manoeuvre apex. If the cost on the body torque represented by  $r_2$  is made smaller, the body torque increases as might be expected. However, the steer torques change very little and the tracking quality is hardly affected. Even quite large body lean torques are apparently ineffective in improving tracking accuracy, although they do affect the rider's leaning relative to the motorcycle. Control actions can be detected when the motorcycle is  $300\text{ m}$  away from the change of direction, by which time the oscillatory steer angle and yaw rate responses are already established. Optimal tracking implies that a high-quality manoeuvre does not start with sudden counter-steering (Cossalter, 2002; Foale, 2002) but rather with a succession of oscillatory inputs of increasing magnitude, designed to excite the weave resonance of the machine. Such a strategy appears to be well aligned with the optimisation objectives, which are accurate path-following with little expenditure of control energy. The rider needs to possess sufficient control bandwidth to make this kind of control possible.



**Figure 32.** Tracking simulation results for a sudden  $15^\circ$  direction change and  $70\text{ m/s}$  speed with  $500\text{ m}$  preview, tight control and equal weights on steer torque and rider lean torque ( $q_1 = 10^4$ ,  $q_2 = 4 * 10^4$ ,  $q_3 = 0$ ,  $r_1 = r_2 = 1$ ). The upper part shows the motorcycle roll angle, the rider upper body relative lean angle, the steer angle and the yaw rate, while the lower part shows the frame twist and the tyre sideslip angles, scaled for convenience.

For the same manoeuvre at the same speed with loose control represented by  $q_1 = 100$ ,  $q_2 = 400$ ,  $q_3 = 0$  and with  $1000\text{ m}$  preview, and again the same weights on steer torque and body torque ( $r_1 = r_2 = 1$ ), the control inputs commence earlier and finish later, and much more corner-cutting occurs. The control inputs are reduced in magnitude markedly, see Figure 34. The peak roll angle is reduced from about  $9^\circ$  to  $4^\circ$ , see Figure 35 and the control inputs and motorcycle path are hardly oscillatory. Effectively, associating higher costs with the use of controls obliges the rider to curtail high-frequency control activity, to spread the control inputs over an extended period of time, and to use smaller torques. Steer torque is lower by about a factor of 4, while the body torques are lower by a factor of about 80. Tyre sideforces and aligning moments are smaller by a factor of about 2, see Figure 36.

Simulations of other manoeuvres at other speeds show the same kind of patterns. If the motions demanded by the tracking task require only low frequency control inputs, then almost perfect tracking performance results. As the path demands become greater, tracking accuracy is sacrificed to an extent that depends on the weights in the cost function. Saving control power is achieved by extending the control inputs in time and using lower magnitudes of control torque. The rider uses more preview and needs to be able to see the road at a greater distance ahead. With limited preview, the rider is

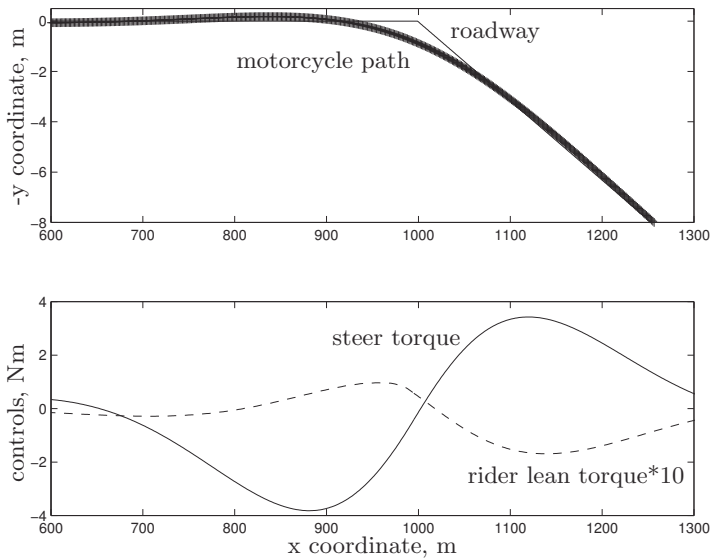


**Figure 33.** Tracking simulation results for a sudden  $15^\circ$  direction change and  $70\text{ m/s}$  speed with  $500\text{ m}$  preview, tight control and equal weights on steer torque and rider lean torque ( $q_1 = 10^4$ ,  $q_2 = 4 * 10^4$ ,  $q_3 = 0$ ,  $r_1 = r_2 = 1$ ). The tracking error and the tyre sideforces and aligning moments are shown, scaled for convenience. The aligning moments are multiplied by 20; the rear one is shown dash-dotted, while the front one is shown dotted.

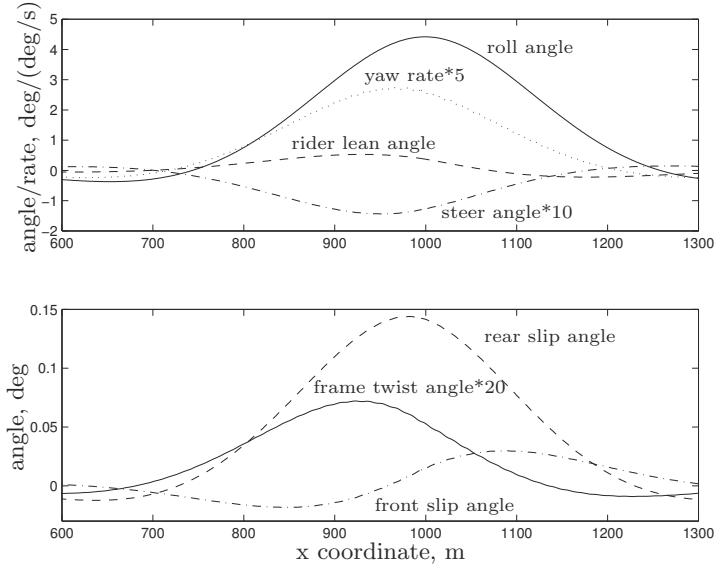
obliged to use tight control, which takes great effort.

## 6 Conclusions

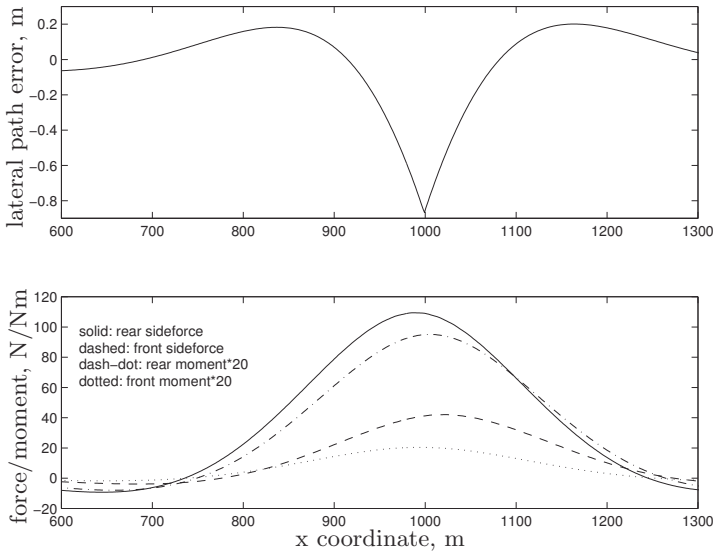
Much of motorcycle handling behaviour can be successfully simulated and to that extent, it can be claimed to be understood. The historic model of Whipple (1899), together with the extensive checking recounted in Schwab et al. (2004), Schwab et al. (2005b) and Schwab et al. (2005a), allows analysis that yields considerable insight into the low-speed, low-frequency behaviour of the two-wheeler, especially the balancing problem. Increasing the trail causes auto-stabilisation to occur at higher speed and lowers the peak divergence rate associated with capsizing. To properly represent a wider range of behaviour over a greater speed range, numerical methods must be relied on. Tyres must be regarded as force generators in response to attitude variables and motions, like camber and sideslip, the forces need to be lagged in relation to slip, frame compliance needs to be accounted for and the rider's lack of rigidity has to be considered, even when confining attention to straight-running. Motorcycles with reasonable speed capability are oscillatory by nature, with weave and wobble constituting distinct modes, with typical frequencies around  $3\text{ Hz}$  and  $8\text{ Hz}$  respectively. A heavy topbox destabilises the high-speed weave significantly.



**Figure 34.** Tracking simulation results for a sudden  $15^\circ$  direction change and  $70\text{ m/s}$  speed with  $1000\text{ m}$  preview, loose control and equal weights on steer torque and rider lean torque ( $q_1 = 100$ ,  $q_2 = 400$ ,  $q_3 = 0$ ,  $r_1 = r_2 = 1$ ).



**Figure 35.** Tracking simulation results for a sudden  $15^\circ$  direction change and  $70\text{ m/s}$  speed with  $1000\text{ m}$  preview, loose control and equal weights on steer torque and rider lean torque ( $q_1 = 100$ ,  $q_2 = 400$ ,  $q_3 = 0$ ,  $r_1 = r_2 = 1$ ).



**Figure 36.** Tracking simulation results for a sudden  $15^\circ$  direction change and  $70\text{ m/s}$  speed with  $1000\text{ m}$  preview, loose control and equal weights on steer torque and rider lean torque ( $q_1 = 100$ ,  $q_2 = 400$ ,  $q_3 = 0$ ,  $r_1 = r_2 = 1$ ). Tracking errors and tyre sideforces and aligning moments are shown, scaled for convenience. The aligning moments are multiplied by 20; the rear one is shown dash-dotted, while the front one is shown dotted.

Front wheel locking is disastrous while rear wheel locking is problematic. Motorcycles have much to gain from effective anti-lock braking systems. Steering dampers are invariably effective in stabilising the wobble mode but they are damaging to the weave mode. Tuned and damped absorbers or other compensator networks have interesting possibilities to maintain the good influence of the steering damper, without having to suffer the negative side. Such devices are not yet fully developed or tested. Active steering compensators may give enhanced benefits, with adaptation of properties to operating conditions being feasible.

Modelling the cornering machine effectively demands a much greater level of detail to be represented and automated multibody modelling software becomes essential to accuracy. The suspension system becomes important and the tyre behaviour becomes complex, needing much data to provide a basis for a wide-ranging representation. A virtual machine and rider can be constructed and used for many purposes. A natural and significant division is into open-loop and closed-loop manoeuvring. Responses to control inputs can be calculated without difficulty but conversion of the results into quality judgments may be difficult and may be very important. Some mysteries remain and further research is needed here. Path-based manoeuvres can be completed only with the aid of a rider control model. Path preview is a key issue and properly incorporating the rider's use of preview is thought to be essential to obtaining good virtual rider tracking



performance.

The application of optimal linear preview control to the problem of accurate path following has been demonstrated and results contain many features that appear descriptive of the real problem. Preview control involves diminishing returns, such that a certain amount is enough for any given set of running conditions. The preview necessary for a motorcyclist is much greater than that needed by a car driver, for a given speed. Body lean control is completely secondary to steer torque control. The optimal steering control is almost unaffected by changes in the use of body control and the tracking quality is hardly influenced either. The theoretical rider considered so far has no bandwidth limitations, so that further work is needed to improve the situation. The way in which the machine design affects the rider's steering control task is also of strong interest and is being studied currently.

## Bibliography

- B. D. O. Anderson and J. B. Moore. *Linear Optimal Control*. Prentice Hall, Englewood Cliffs, NJ, 1971.
- A. Aoki. Experimental study on motorcycle steering performance. *SAE 790265*, 1979.
- B. Bayer. Flattern und pendeln bei krafträdern. *Automobil Industrie*, 2:193–197, 1988.
- F. Biral, D. Bortoluzzi, V. Cossalter, and M. Da Lio. Experimental study of motorcycle transfer functions for evaluating handling. *Vehicle System Dynamics*, 39(1):1–26, 2003.
- A. Clerx. Stijfheid en sterkte van motorfietsframes. Technical report, Department of Mechanical Engineering, Technische Hogeschool Eindhoven, 1977.
- V. Cossalter. *Motorcycle Dynamics*. Race Dynamics, Greendale (WI), 2002.
- E. Döhning. Steering wobble in single-track vehicles. *ATZ*, 58(10):282–286, 1956.
- Dunlop Tyre Company. Wobble and weave. c1977.
- D. J. Eaton. Lateral dynamics of the uncontrolled motorcycle. In *2<sup>nd</sup> Int. Congress on Automotive Safety*, San Fransisco, 1973.
- S. Evangelou, D. J. N. Limebeer, R. S. Sharp, and M. C. Smith. Steering compensation for high-performance motorcycles. In *Proceedings of 43rd Control Decision Conference*, pages 749–754, Atlantis, Paradise Island, Bahamas, December 2004.
- S. Evangelou, D. J. N. Limebeer, R. S. Sharp, and M. C. Smith. Control of motorcycle steering instabilities: passive mechanical compensators incorporating inerters. *IEEE Control Systems Magazine*, 26(5):78–88, October 2006a.
- S. Evangelou, D. J. N. Limebeer, R. S. Sharp, and M. C. Smith. An h-inf loop-shaping approach to steering control for high-performance motorcycles. In Bruce A. Francis, Malcolm C. Smith, and Jan. C. Willems, editors, *Control of Uncertain Systems: Modelling, Approximation, and Design*, pages 257–275, Berlin Heidelberg, 2006b. Springer-Verlag.
- S. Evangelou, D. J. N. Limebeer, R. S. Sharp, and M. C. Smith. Mechanical steering compensators for high-performance motorcycles. *ASME Journal of Applied Mechanics*, 2006c. in press.

- T. Foale. *Motorcycle handling and chassis design – the art and science*. Tony Foale Designs, Spain, 2002.
- R. Frezza and A. Beghi. Simulating a motorcycle rider. In W. Kang, editor, *New Trends in Nonlinear Dynamics and Control*, Berlin/Heidelberg, 2003. Springer Verlag.
- C. G. Giles and R. S. Sharp. Static and dynamic stiffness and deflection mode measurements on a motorcycle, with particular reference to steering behaviour. In *Proc. Inst. Mech. Eng./MIRA Conference on Road Vehicle Handling*, pages 185–192, London, 1983. Mechanical Engineering Publications.
- K. Gurney. *Introduction to Neural Networks*. Routledge, New York, 1997.
- F. D. Hales. Lateral stability problems of simply articulated vehicles. In *Proceedings of IUTAM Symposium on Recent Progress in Linear Mechanical Vibrations*, pages 17–34, 1965.
- G. Jennings. A study of motorcycle suspension damping characteristics. *SAE 740628*, 1974.
- C. Juden. Cycling science. *Cycletouring*, April/May:153–153, 1987.
- T. Katayama, A. Aoki, and T. Nishimi. Control behaviour of motorcycle riders. *Vehicle System Dynamics*, 17(4):211–229, 1988.
- C. Koenen and H. B. Pacejka. The influence of frame elasticity, simple rider body dynamics, and tyre moments on free vibrations of motorcycles in curves. In *Proc. 7<sup>th</sup> IAVSD Symposium on Dynamics of Vehicles on Roads and on Railway Tracks*, pages 53–65, Lisse, 1982. Swets and Zeitlinger.
- D. J. N. Limebeer, S. Evangelou, and R. S. Sharp. Stability of motorcycles under acceleration and braking. In *Proc. DETC’2001, 18th Biennial Conference on Mechanical Vibration and Noise*, Pittsburgh, PA, 2001a. ASME.
- D. J. N. Limebeer and R. S. Sharp. Bicycles, motorcycles and models: Single-track vehicle modeling and control. *IEEE Control Systems Magazine*, 26(5):34–61, October 2006.
- D. J. N. Limebeer, R. S. Sharp, and S. Evangelou. The stability of motorcycles under acceleration and braking. *Proc. I MECH E Part C Jour. Mech. Eng. Sci.*, 215(c9): 1095–1109, 2001b.
- D. J. N. Limebeer, R. S. Sharp, and S. Evangelou. Motorcycle steering oscillations due to road profiling. *Transactions of the ASME, Journal of Applied Mechanics*, 69(6): 724–739, 2002.
- N. Louam, D. A. Wilson, and R. S. Sharp. Optimal control of a vehicle suspension incorporating the time delay between front and rear wheel inputs. *Vehicle System Dynamics*, 17(6):317–336, 1988.
- J. S. McKibben. Motorcycle dynamics – fact, fiction and folklore. *SAE 780309*, 1978.
- J. P. Meijaard and A. A. Popov. Application of non-linear dynamics to instability phenomena in motorcycles. In M. Abe, editor, *Proc. 18<sup>th</sup> IAVSD Symposium on Dynamics of Vehicles on Roads and on Tracks*, pages 182–191, 2004. Suppl. Vehicle System Dynamics.
- J. P. Meijaard and A. A. Popov. Numerical continuation of solutions and bifurcation analysis in multibody systems applied to motorcycle dynamics. *Nonlinear Dynamics*, 43(1-2):97–116, 2006.

- P. C. Mitiguy and K. J. Reckdahl. The definition of product of inertia. *Working Model Technical Paper*, 1998.
- T. Nishimi, A. Aoki, and T. Katayama. Analysis of straight running stability of motorcycles. In *10<sup>th</sup> International Technical Conference on Experimental Safety Vehicles*, pages 1080–1094, Oxford, July 1-5 1985.
- W. M. Otto. Effect of motorcycle accessories on stability. In *Proc. International Motorcycle Safety Conference*, pages 1560–1581, Washington DC, May 1980. Motorcycle Safety Foundation, Linthicum MD 21090.
- H. B. Pacejka. *Tyre and Vehicle Dynamics*. Butterworth Heinemann, Oxford, 2002. ISBN 0-7506-5141-5.
- G. Prokop and R. S. Sharp. Performance enhancement of limited bandwidth active automotive suspensions by road preview. *IEE Proceedings Control Theory and Applications*, 142(2):140–148, 1995.
- R. S. Rice. Rider skill influences on motorcycle maneuvering. In *Motorcycle Dynamics and Rider Control, SP-428*, pages 79–90, Warrendale PA, 1978. SAE Inc.
- G. E. Roe and T. E. Thorpe. A solution of the low-speed wheel flutter instability in motorcycles. *Proc. I MECH E Part C Jour. Mech. Eng. Sci.*, 18(2):57–65, 1976.
- A. L. Schwab, J. P. Meijaard, and J. D. G. Kooijman. Experimental validation of a model of an uncontrolled bicycle. In C. A. Mota Soares et al., editor, *Proc. III European Conference on Computational Mechanics: Solids, Structures and Coupled Problems in Engineering*, pages 1–16, 2006.
- A. L. Schwab, J. P. Meijaard, and J. M. Papadopoulos. Benchmark results on the linearized equations of motion of an uncontrolled bicycle. In *Proc. Second Asian Conference on Multibody Dynamics*, pages 1–9, 2004.
- A. L. Schwab, J. P. Meijaard, and J. M. Papadopoulos. Benchmark results on the linearized equations of motion of an uncontrolled bicycle. *KSME International Journal of Mechanical Science and Technology*, 19(1):292–304, 2005a.
- A. L. Schwab, J. P. Meijaard, and J. M. Papadopoulos. A multibody dynamics benchmark on the equations of motion of an uncontrolled bicycle. In *Proc. ENOC-2005, Eindhoven*, pages 511–521, 2005b.
- A. Shaeri, D. J. N. Limebeer, and R. S. Sharp. Nonlinear steering oscillations of motorcycles. In *Proc. 43rd IEEE Conference on Decision and Control*, pages 773–778, Atlantis, Paradise Island, Bahamas, 14-17 December 2004.
- R. S. Sharp. The stability and control of motorcycles. *Proc. I MECH E Part C Jour. Mech. Eng. Sci.*, 13(5):316–329, 1971.
- R. S. Sharp. The stability of motorcycles in acceleration and deceleration. In *Inst. Mech. Eng. Conference Proceedings on “Braking of Road Vehicles”*, pages 45–50, London, 1976. MEP.
- R. S. Sharp. A review of motorcycle steering behavior and straight line stability. In *Motorcycle Dynamics and Rider Control, SP-428*, pages 1–6, Warrendale PA, 1978. SAE Inc.
- R. S. Sharp. The lateral dynamics of motorcycles and bicycles. *Vehicle System Dynamics*, 14(4-6):265–283, 1985.

- R. S. Sharp. Vibrational modes of motorcycles and their design parameter sensitivities. In *Vehicle NVH and Refinement*, pages 107–121, London, 1994. Mech. Eng. Publications.
- R. S. Sharp. Some contemporary problems in road vehicle dynamics. *Proc. I MECH E Part C Jour. Mech. Eng. Sci.*, 214(1):137–148, 2000.
- R. S. Sharp. Stability, control and steering responses of motorcycles. *Vehicle System Dynamics*, 35(4–5):291–318, 2001.
- R. S. Sharp. Driver steering control and a new perspective on car handling qualities. *Journal of Mechanical Engineering Science*, 219(C8):1041–1051, 2005.
- R. S. Sharp. Optimal linear time-invariant preview steering control for motorcycles. In S. Bruni and G. R. M. Mastinu, editors, *The Dynamics of Vehicles on Roads and on Railway Tracks*, page in press, London, 2006. Taylor and Francis. Supplement to Vehicle System Dynamics, Vol. 44.
- R. S. Sharp. Motorcycle steering control by road preview. *Trans. ASME Journal of Dynamic Systems, Measurement and Control*, page in press, 2007.
- R. S. Sharp and C. J. Alstead. The influence of structural flexibilities on the straight running stability of motorcycles. *Vehicle System Dynamics*, 9(6):327–357, 1980.
- R. S. Sharp, S. Evangelou, and D. J. N. Limebeer. Advances in the modelling of motorcycle dynamics. *Multibody System Dynamics*, 12(3):251–283, 2004.
- R. S. Sharp, S. Evangelou, and D. J. N. Limebeer. Multibody aspects of motorcycle modelling with special reference to autosim. In J. A. C. Ambrosio, editor, *Advances in Computational Multibody Systems*, pages 45–68, Dordrecht, The Netherlands, 2005a. Springer-Verlag.
- R. S. Sharp, D. J. N. Limebeer, and A. Shaeri. Nonlinear vibrations of motorcycles forced by regular road roughness. In *Proc. ENOC-2005, Eindhoven*, pages 2226–2234, 2005b.
- R. S. Sharp and V. Valtetsiotis. Optimal preview car steering control. In P. Lugner and K. Hedrick, editors, *ICTAM Selected Papers from 20th International Congress*, pages 101–117, Lisse, 2001. Swets and Zeitlinger. supplement to Vehicle System Dynamics Vol. 35.
- M. C. Smith. Synthesis of mechanical networks: The inerter. *IEEE Transactions on Automatic Control*, 47(10):1648–1662, October 2002.
- P. T. J. Spierings. The effects of lateral front fork flexibility on the vibrational modes of straight-running single-track vehicles. *Vehicle System Dynamics*, 10(1):21–35, 1981.
- M. Sugizaki and A. Hasegawa. Experimental analysis of transient response of motorcycle rider systems. *SAE 881783*, 1988.
- T. Takahashi, T. Yamada, and T. Nakamura. Experimental and theoretical study of the influence of tires on straight-running motorcycle weave response. *SAE 840248*, 1984.
- B. Thomson and H. Rathgeber. Automated systems used for rapid and flexible generation of system models exemplified by a verified passenger car and a motorcycle model. In J. Karl Hedrick, editor, *The Dynamics of Vehicles on Roads and on Railway Tracks*, pages 645–654, Lisse, 1984. Swets and Zeitlinger.
- M. Tomizuka. Optimal linear preview control with application to vehicle suspension - revisited. *ASME Journal of Dynamic Systems, Measurement, and Control*, 98:309–315, 1976.

- M. K. Verma, R. A. Scott, and L. Segel. Effect of frame compliance on the lateral dynamics of motorcycles. *Vehicle System Dynamics*, 9(3):181–206, 1980.
- Y. Watanabe and K. Yoshida. Motorcycle handling performance for obstacle avoidance. In *2<sup>nd</sup> Int. Congress on Automotive Safety*, San Fransisco, 1973.
- D. H. Weir and J. W. Zellner. Lateral-directional motorcycle dynamics and rider control. In *Motorcycle Dynamics and Rider Control, SP-428*, pages 7–31, Warrendale PA, 1978. SAE Inc.
- D. H. Weir and J. W. Zellner. Experimental investigation of the transient behaviour of motorcycles. *SAE 790266*, 1979.
- F. J. W. Whipple. The stability of the motion of a bicycle. *Quarterly Journal of Pure and Applied Mathematics*, 30(120):312–321, 1899.
- J. W. Zellner and D. H. Weir. Development of handling test procedures for motorcycles. In *Motorcycle Dynamics and Rider Control, SP-428*, pages 91–100, Warrendale PA, 1978. SAE Inc.

# Control of Rail Vehicles: Mechatronic Technologies

Roger Goodall

Department of Electronic and Electrical Engineering,  
Loughborough University, Leicestershire, UK

**Abstract** Mechatronic technologies for railway vehicles improve their dynamical performance allowing higher speeds, they result in more comfort for the passengers and less wear of the tracks. Thus, transportation by rail is getting more attractive and less expensive. With reference to mechatronically-designed rail vehicle six projects completed by the author during the last decade are summarized.

## 1 Introduction

Conventional rail vehicles consist of purely mechanical parts for guidance and suspension. In particular, the rigid wheelset is subject to strong wear on curved tracks, and a spring-damper suspension does not enable the wagon for tilting to compensate the centrifugal forces impacting the passengers laterally. Many new features are offered by mechatronic components added to a rail vehicle. Six corresponding projects are summarized in this contribution. For more detailed information references are given, and the reader is encouraged to make use of them.

Mechatronics is where computing, electronics and control are brought together in the design of a mechanical system in order to produce a mechatronic solution. For a railway vehicle, there are a number of benefits that may arise from the use of mechatronic's: firstly, it may simply make the product cheaper, or may better although what better is needs defining. Perhaps it will be possible for the vehicle to be lighter, and even perhaps faster. All of these possible benefits will in one way or another have some sort of business benefit that can be exploited for the benefit of the railway as a whole.

Of course all of the above-mentioned benefits, although they can be quantified in a business sense, they are essentially non-technical in nature. Therefore it is possible to be more specific about the benefits that can arise. The primary possibilities in a technical sense are as follows:

- improved dynamic performance, e.g. higher speed,
- lighter, mechanically-simpler vehicles,
- track-friendly vehicles,
- reduced cost of operation, in particular lower track maintenance costs.

A number of these specific technical benefits are be considered in the paragraphs below when the particular opportunities for applying mechatronics are looked at in more detail, but first it's useful to briefly review the history of railway vehicle design.

### **1825-1960s: Empirical Design Period**

For the first 140 years of railway vehicle design the vehicles were largely designed from an empirical point of view. That is not to say that no calculations were carried out, but it is certainly true that, from the point of view of the vehicle dynamics system as a whole, the ability to understand, model and calculate the overall dynamic performance was extremely limited - hence the statement that developments for many, many years were essentially empirical.

### **1960s-1990s: Analytical Design Period**

In the 1960s the science of railway vehicle dynamics was born: various scientific threads were brought together in order to produce a complete theoretical understanding of the way that railway wheelsets worked, and this was combined with the availability of significant computing power which meant that it was possible to predict the performance of a complete railway vehicle, something that had previously not been possible. This heralded the arrival of what can be called an analytical design period which lasted from the 1960s through until the 1990s.

### **1990s- ???: Mechatronic Design Period**

During the 1990s, in other industries such as the aircraft and automotive industries, the power that became available from designing the mechanical system in conjunction with the electronics, computing and control, i.e. the use of mechatronics, was realised in a variety of research and development programmes. Although the railway industry is, perhaps naturally, somewhat behind these other two industries, nevertheless a variety of developments are being considered currently which imply that a "mechatronic design period" is close to happening for railway vehicles.

## **2 Mechatronic Technologies for Trains of the Future**

The primary motivation and starting point for research project was the assumption that future rail vehicles will have sensors, controllers and actuators affecting the basic operation of their running gear. The project was a "Basic Research" project led by Loughborough University and involving six other European organisations. A practical engineering focus that recognised commercial requirements was taken from the outset, the overall aim being to understand and develop the engineering science applicable to mechatronics for rail vehicles, and to assess the possible impacts on rail vehicle design. Specification objectives were as follows:

1. To develop a fundamental understanding of the dynamic response of lightweight rail vehicles with active controls.
2. To develop methods of analysis for advanced vehicles emphasizing configurations which would take full advantage of emerging control technology (leading to reduced weight, lower cost, lower car-body structural vibrations, etc.).
3. To identify systems architectures (sensors, actuators, processing) which would provide the level of safety, reliability, and maintainability needed for an operational railway.

More detail are provided by Ellis and Goodall (1999) and Goodall and Kortum (2002).

### **3 Active Steering and Stability Control for a Railway Bogie**

This project is fully-funded by Bombardier Transportation, and is concerned with translating some of the ideas that have been developed during previous research studies into practice. Bombardier have modified a conventional railway bogie so that each axle can be separately controlled using electrical actuators, and the research objectives are to produce practical control strategies that both steer the bogie properly through curves and provide stable operation for high speed running. Active stabilisation of the prototype bogie at speeds over 300 Km/h has been demonstrated in tests on a "roller rig" in Munich. More detail are provided by Goodall and Li (2000) and Pearson et al. (2004).

### **4 Dynamics and Control Study of Independently-driven Railway Wheels**

This project is funded by the UK research council EPSRC and carried out in collaboration with Stored Energy Technology Ltd, a small company based in Derby. It is concerned with developing control strategies for a novel form of rail vehicle in which traction motors integrated within the wheels are independently controlled so that the vehicle is not only driven and braked, but also guided around curves. The challenge is to devise a controller for what is a dynamically-complex system, and develop this into a practical, working system for a demonstrator vehicle. More detail are provided by Mei et al. (2002).

### **5 High-integrity Sensing for Active Steering of Railway Vehicles**

Active steering of both conventional solid-axle and independently-rotating wheelsets for railway vehicles offers a number of advantages for vehicle design, but remains a highly speculative area of research. This project investigated safety-critical measuring systems for active steering of railway vehicles, because the provision of high integrity measurements is an important practical problem for implementing any of the control strategies which are possible. In particular, some of the natural choices for feedback such as lateral wheel/rail displacement are very difficult to measure in a cost-effective manner, and so the primary aim was to determine practical techniques for estimating such variables using wheelset measurements that can not only be implemented cheaply and easily but would also offer the potential for high integrity. More detail are provided by Li and Goodall (2000).

### **6 Control Strategies for Flexible-bodied Railway Vehicles**

Dealing with body flexibility is an increasing difficulty in suspension design on modern high-speed rail vehicles and effective active strategies for dealing with the problem may enable lighter more flexible bodies to be used. The project was an advanced active



railway suspension study that investigated ways of applying classical and optimal control strategies based upon the idea of using skyhook damping to actively control the flexible body modes.

The study included the novel idea of adding a third actuator at the centre of the vehicle body to suppress primarily the flexible body modes (i.e. first symmetrical and first asymmetrical) of the railway vehicle body. This was used together with two actuators more conventionally located across the front and rear secondary suspensions. Classical and optimal control strategies for both two and three actuators options were investigated, and the impact of the actuators' dynamic characteristics included using electro-hydraulic actuators for the front and rear actuators and an electromagnetic actuator for the centre. More detail are provided by Foo and Goodall (2000).

## 7 Advanced Control Strategies for Tilting Trains

The use of tilting bodies on railway vehicles is becoming increasingly widespread: a number of well-established services using tilt technology already exist around the world, and tilting trains are appearing again in the UK. The motivation for tilting railway vehicles is that they give a cost-effective means of achieving a substantial reduction in journey time by increasing the vehicle speed during curves.

Early tilt controller designs were based upon local vehicle measurements, however at that time this approach did not prove very successful. Nowadays most European manufacturers use the so called 'precedence' control scheme, utilising measurements from precedent vehicles to achieve 'precedence' information. However, achieving a satisfactory tilt control strategy using local vehicle measurements only is an important research target because of the system simplifications and more straightforward failure detection.

The work achieved a comprehensive study of tilt control, and in particular investigated advanced control techniques with the particular objective of identifying effective strategies which can be applied to each vehicle independently, i.e without using precedence control. More detail are provided by Zolotas et al. (2007).

## 8 Conclusion

In this contribution six projects on control of rail vehicles are summarized showing the high engineering and economics potential of mechatronic technologies. In particular, tilting trains require advanced control strategies resulting in higher speeds on existing tracks.

## Bibliography

- B. Ellis and R.M. Goodall. The Mechatronic Train: Requirements and Concepts. In *Proceedings of World Congress on Railway Research*, pages 1–7, Tokyo, Japan, October 1999. Railway Technical Research Institute.
- T.H.E. Foo and R.M. Goodall. Active Suspension Control of Flexible-bodied Railway Vehicles Using Electro-hydraulic and Electro-magnetic Actuators. *Control Engineering Practice*, 8:507–518, 2000.

- 
- R.M. Goodall and W. Kortum. Mechatronic Developments for Railway Vehicles of the Future. *Control Engineering Practice*, 10(8):887–898, 2002.
- R.M. Goodall and H. Li. Solid Axle and Independently-Rotating Railway Wheelsets - A Control Engineering Assessment of Stability. *Vehicle System Dynamics*, 33:57–67, 2000.
- H. Li and R.M. Goodall. Fault-Tolerant Sensing for Rail Vehicles. In *Proceedings of 14th International Conference on Systems Engineering ICSE 2000*, pages 381–385, University of Coventry, September 2000.
- T. X. Mei, H. Li, R. M. Goodall, and A. H. Wickens. Dynamics and control assessment of rail vehicles using permanent magnet wheel motors. *Vehicle System Dynamics*, 37(supplement):326–337, 2002.
- J.T. Pearson, R.M. Goodall, T.X. Mei, and G. Himmelstein. Active stability control strategies for a high speed bogie. *Control Engineering Practice*, 12:1381–1391, 2004.
- A.C. Zolotas, R.M. Goodall, and G. Halikias. Recent Results in Tilt Control Design and Assessment of High Speed Railway Vehicles. *IMECHE Journal of Railway and Rapid Transit*, 2007.

# RailCab System: Engineering Aspects

Joachim Lückel <sup>\*</sup>, Horst Grotstollen <sup>†</sup>, Markus Henke <sup>\*\*</sup> Thorsten Hestermeyer <sup>‡</sup>,  
Xiaobo Liu-Henke <sup>‡</sup>

<sup>\*</sup> Mechatronics & Control, University of Paderborn, Germany

<sup>†</sup> Power Electronics & Electrical Drive Systems, University of Paderborn, Germany

<sup>\*\*</sup> formerly: Power Electronics & Electrical Drive Systems, University of Paderborn, Germany

<sup>‡</sup> formerly: Mechatronics & Control, University of Paderborn, Germany

**Abstract** The test vehicles of the Paderborn RailCab system, especially developed and scaled down for testing purposes, are conceived as complex mechatronic structures. In designing their components for driving, energy management, suspension system, control and communication system we broke fairly new ground. As university research teams have only extremely limited personal and financial resources at their disposal, realisation of such a system was only possible using new design methods particularly customised to mechatronics. These comprise a model-based approach with appropriate modular-hierarchical structuring for a holistic observation that includes primarily the elements for control and information processing from the outset. In view of the drive and suspension components dealt with in detail in this paper, this purely computer-aided approach was complemented by extensive lab tests in hardware-in-the-loop technology before the test vehicles were actually constructed and tested.

## 1 Introduction to the RailCab System

Almost ten years ago a group of professors in engineering at the University of Paderborn jointly started up a concept, along with its technical basics, of a novel traffic system that was to translate the individual possibilities of the motorcar to the well-worn railway tracks. Such autonomous traffic systems have been in the discussion for many years, especially in the English-speaking world, and labelled “personal rapid transport” (PRT) or “personal automated transport” (PAT); some of them have even been realised as prototypes. The particularity of the Paderborn approach is the combination and fine-tuning of novel and well-proven technical components for the drive, chassis, and steering/control technology along with new modular-hierarchical elements of a distributed global/local information processing. The focus was and still is on easing resp. diminishing traffic in densely populated urban regions (e.g. the Ruhr area) and/or connecting and integrating sparsely populated areas (e.g. the region of East-Westphalia and Lippe). This requires solving the conflict between individual and mass transport.

The essential features of such systems are the following:

- many small vehicles,
- driverless, fully automated operation,
- uninterrupted journey on demand.

As has been said, such PRT/PAT systems have been in the discussion for more than thirty years. However, a successful realisation depends essentially on the availability of viable technical solutions for all components - a condition that has only been fulfilled for a few years, especially regarding the necessary information technology.

The elements of the Paderborn design are the following:

- a modular-hierarchical construction-kit structure ensuring an exceptionally high flexibility, scalability, and adaptability as regards technical and application-specific features, at a speed spectrum of 80 to 160 km/h, to be applied both in passenger and freight transport on short and long distances,
- using existing rails and tracks with high reliability, disposability, and safety,
- using the most recent technologies of drive technology and energy management, mechatronic chassis technology, distributed global/local information processing for control and monitoring technology and a novel globally/locally distributed logistics.

The developments in the RailCab system pertain to the vehicles, the vehicle system, the convoys, and the superordinated traffic system.

As to the origin of the RailCab system:



Figure 1: Conventional railway.

The conventional railway is marked by

- a proven technology,
- an extensive route network,
- centralistic, monolithic features.



Figure 2: The Transrapid.

The relatively new Transrapid system is marked by

- an innovative drive- and supporting/guiding technology,
- a high top speed,
- but also centralistic, monolithic features,
- requiring an entirely new route network.

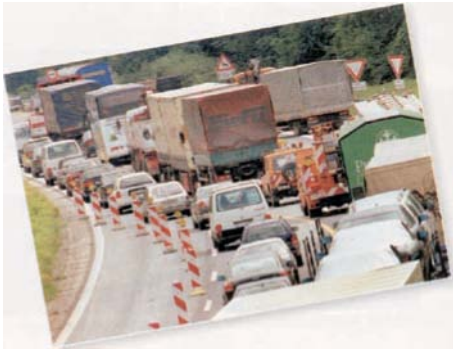


Figure 3: Individual traffic.

The individual traffic system “motorcar” is marked by

- high individual flexibility and adaptability,
- high comfort,
- but an extensive use of the resulting leeway has soon gone haywire and thus brought about an overloaded road network.



Figure 4: RailCab system.

The RailCab system is marked by

- high flexibility and high individual comfort,
- high average speed,
- decentralised, distributed features,
- innovative mechatronic system technology.

### 1.1 The RailCab Vehicle, the Drive

At the outset the Paderborn group primarily and intensively discussed a suitable drive principle, with one proposal being the diesel engine, well proven in motorcars and advanced in many years’ work. However it entails problems as regards reliability and availability. A comparison of the signal flows of diesel, conventional electrical drive, and linear motor shows substantial differences as schematically displayed in Figures 5-7.

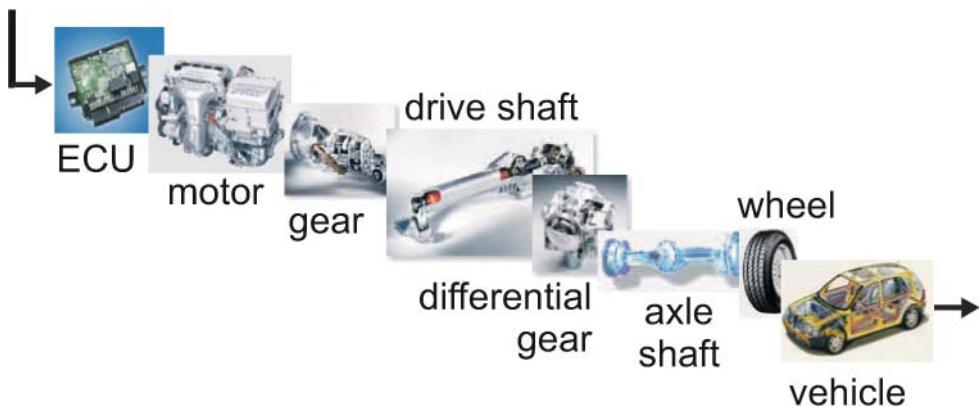


Figure 5: Signal flow of a diesel engine motor car.

The reliability is a result of the multiplikation product made up of the single probabilities; with many aggegrates this means a considerable change for the worse.

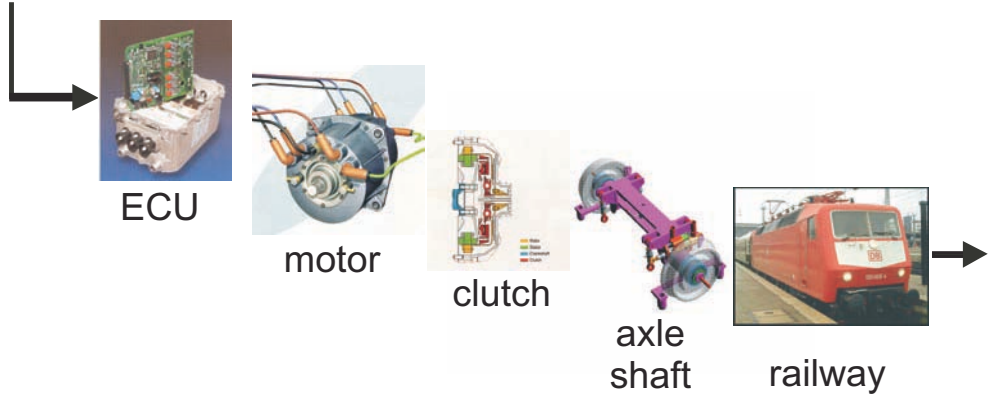


Figure 6: Signal flow of a railway drive.

In the conventional electrical drive in locomotives the signal flow is shorter and thus more reliable, as Figure 6 reveals.

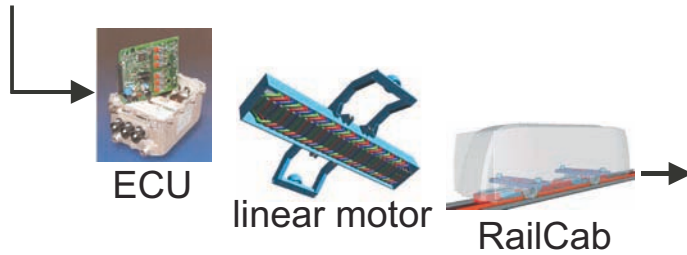


Figure 7: Signal flow of a linear-motor drive.

The shortest direct way is the one of the linear motor; this is why it became the choice for the RailCab.

### 1.2 The Linear Motor

There is no friction, energy is transferred. The linear motors are laid out as double-fed asynchronous motors. In this way several vehicles can run on one stator segment individually and thus make up convoys. The motor contains no movable parts; it is therefore contactless, and there is no need for pantographs.



Figure 8: Double-fed linear motors.

The vehicles are able to manage steep slopes and narrow bends, and the motor can be adapted to the topology of the landscape.



Figure 9: Rotor in the shuttle.

### 1.3 The Mechatronic Chassis, the Steering

The vehicles run on single axles with a set of independently rotating wheels. The axle itself is steerable so that the rims are only needed in emergencies; otherwise the control keeps the wheels centered on the rails. Moreover, only passive, fixed switches are employed. The direction is determined by the feedforward parameters and steering at the wheels. Thus a vehicle can exit the convoy at full speed.

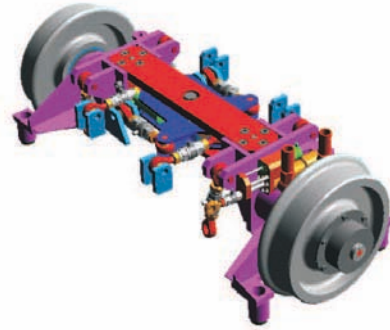


Figure 10: Steerable axle.

### 1.4 The Mechatronic Chassis, the Suspension/Tilt Module

Due to a serial connection of an airspring (for the damping of high frequencies) and three hydraulic actuators each (for the damping of vibrations in the lower frequency range and for the chassis tilt), the suspension/tilt module of the mechatronic chassis provides a noticeable decoupling of vibrations between coach-body and chassis. With the lightweight vehicles and their single axles, the result is high ride comfort combined with careful treatment of humans and material.

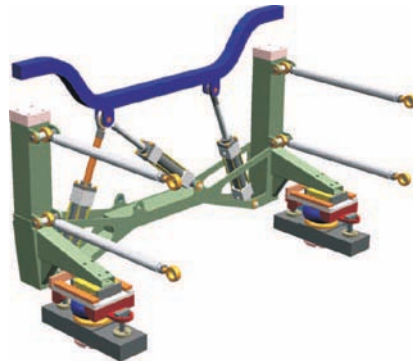


Figure 11: Suspension/tilt module.

### 1.5 The Control Technology and Logistics

The control and logistics system consists of a distributed global as well as local information processing. In it the activating functions affecting the vehicle motions are shifted on board the vehicles in the shape of mobile and local devices while the guiding functions controlling these motions are stationary and global; they are located on different hierarchical levels up to that of logistics (Figure 12). Thus it will be possible to direct all shuttles individually while simultaneously keeping track of the overall system operations.

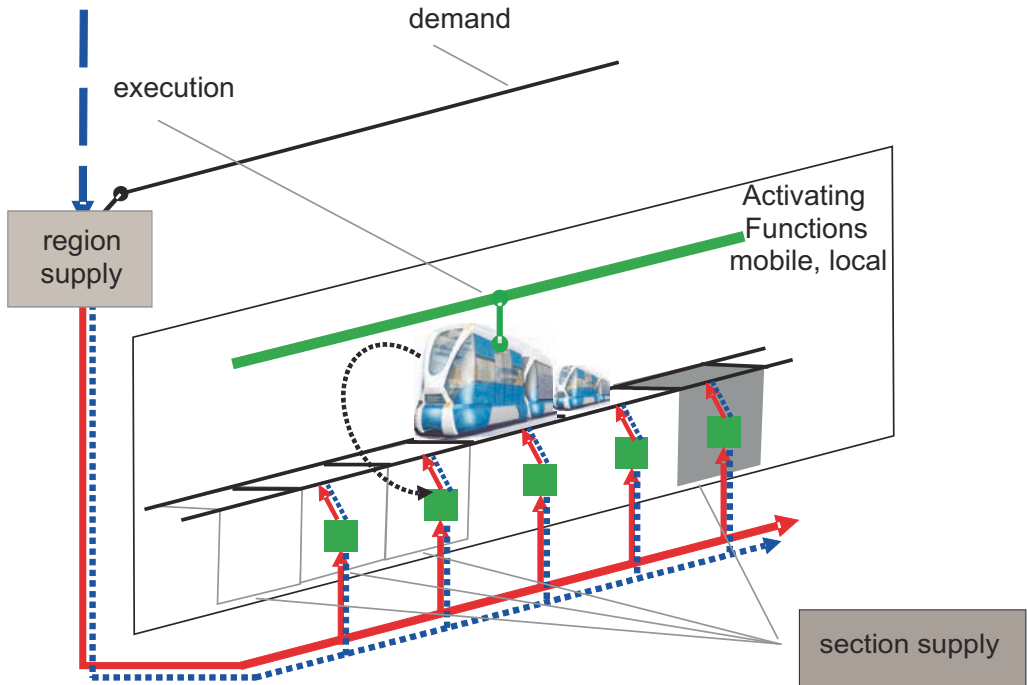


Figure 12: The modular-hierarchical control and logistics system.



## 1.6 Test Track and Test Vehicle in Paderborn

In order to validate the complex mechatronic technology with its tightly interlinked modules a test track was built up close to the University of Paderborn where RailCab vehicles can be tested at the reduced scale of 1:2.5 (Figure 13). The track is 550 m long, has a maximum slope of almost 6% and relatively narrow bends. The vehicles are 3 m long, 1 m high and weigh about 1.2 tons. Their maximum speed is 10 m/sec. The tests fully confirmed the theoretical assumptions. Meanwhile there are two vehicles running simultaneously and independently of one another, the focus being on the convoy tests. Particular attention is paid to steering and track guidance on the passive switch - a concept with a promising future whose potential has already been fully confirmed.

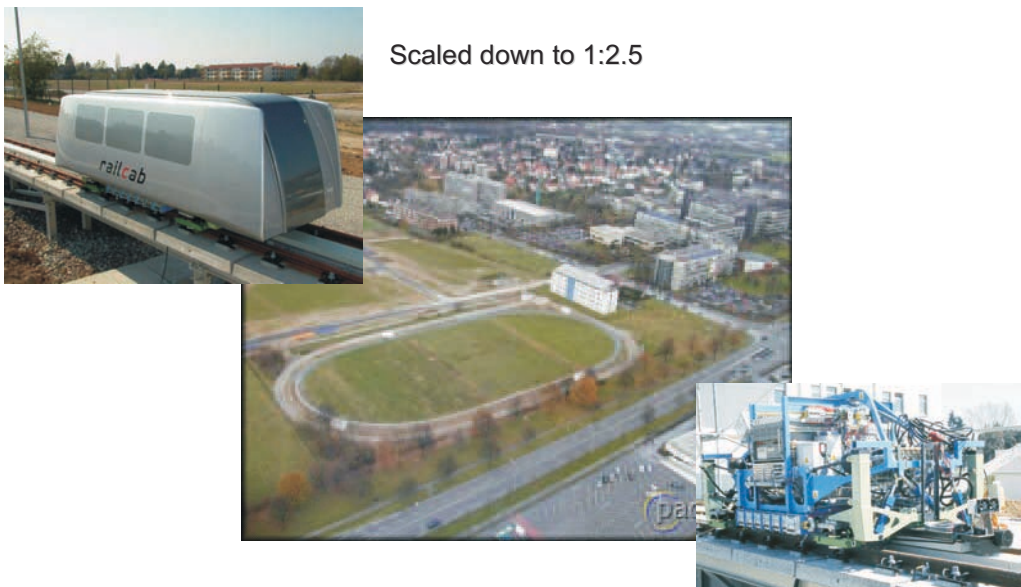


Figure 13: Test vehicle and test track in Paderborn.

### 1.7 The RailCab System, the Convoy

In order to secure sufficient throughput in spite of the vehicles operating individually and on demand, the latter are combined in convoys on the main tracks. Here they run tightly interlocked yet without a mechanical coupling, an arrangement which also provides for a reduced aerodynamical resistance (see Figure 15). The formation and disintegration of such convoys are entirely computer-aided and fully automated.



Figure 14: Formation of a convoy.

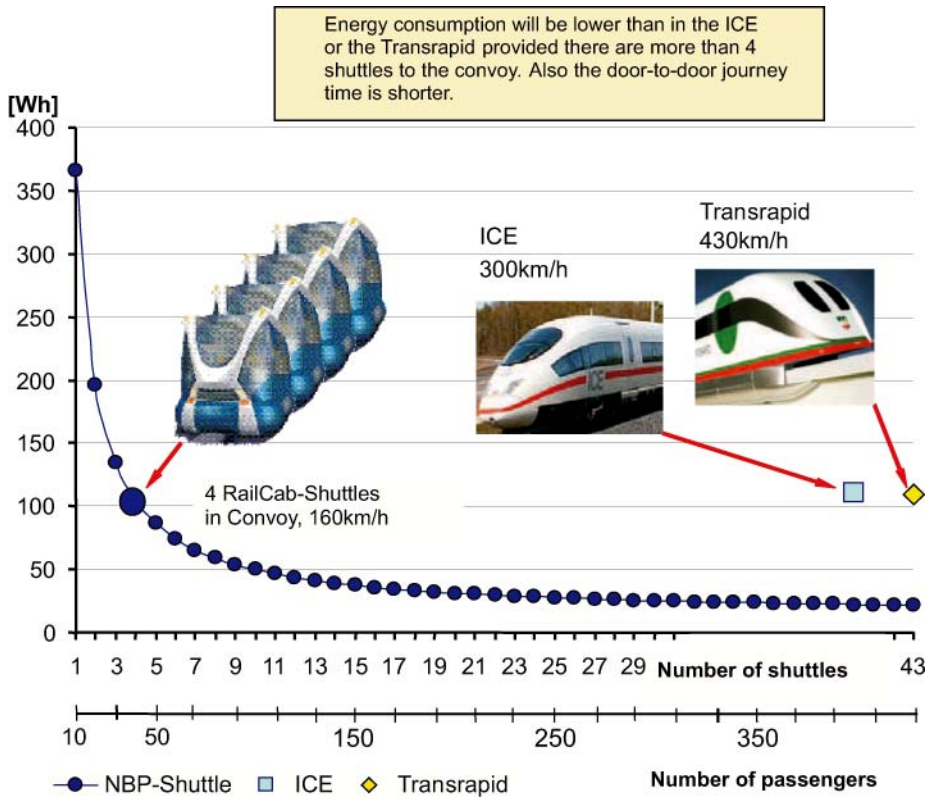


Figure 15: Primary energy consumption per person and km at respective speed.

## 2 Drive and Control Structure of the RailCab

### 2.1 Linear Induction Motor

With regard to energy consumption convoy operation of RailCab vehicles is an important demand Lückel et al. (1999). While forming or splitting a convoy participating vehicles must move with different velocities which is not possible with synchronous motors. Therefore two types of linear induction motor are implemented at the NBP test track, both making use of the same vehicle-side motor part.

**Doubly-Fed Long-Stator Motor [Henke and Grotstollen (2000), Grotstollen (2005), Henke (2003)]** A doubly-fed induction motor exists where the track is equipped with a three-phase long-stator similar to that of synchronous motors used at Transrapid and Maglev, Fig. 16. Due to three-phase excitation both parts of the doubly-fed induction motor generate magnetic traveling waves which can move with regard to both motor

windings. Consequently speed of a vehicle is not fixed to the (synchronous) speed of traveling waves but can be varied individually.

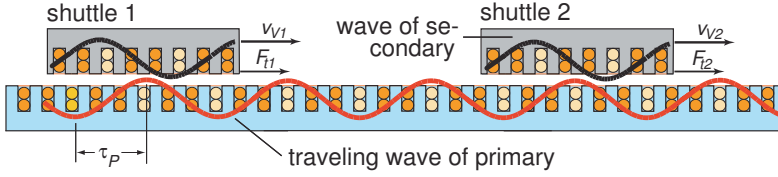


Figure 16: Doubly-fed linear induction motor with long-stator

Thrust force  $F_t$  of the motor is caused by attracting and repelling forces between the traveling waves of primary and secondary. It is calculated from

$$F_t = \frac{3\pi}{2\tau_P} \cdot L_{12} \cdot \Im\{\underline{i}_1 \cdot \underline{i}_2^*\} = \frac{3\pi}{2\tau_P} \cdot L_{12} \cdot i_1 i_2 \cdot \sin \alpha_{1-2}, \quad (2.1)$$

where  $\underline{i}_1$ ,  $\underline{i}_2$  represent the vectors of the primary and secondary current ( $\tau_P$ : pole pitch,  $L_{12}$ : mutual inductance). The spatial phase displacement  $\alpha_{1-2}$  of the traveling waves matches the phase shift of the related current vectors if these are considered in the same coordinate frame. Obviously constant thrust force cannot be achieved but with phase shift  $\alpha_{1-2}$  being constant meaning that an exact synchronization of traveling waves is a must.

When the speed  $v_V$  of a vehicle does not match the (synchronous) speed  $v_1 = 2\tau_P \cdot f_1$ , determined by the traveling waves, voltage is induced to the secondary. This is why energy can be exchanged between primary and secondary winding as at any induction motor. For the power  $P_B$  transferred from the primary to the secondary winding the following equation holds,

$$P_B = 3\pi \cdot f_2 \cdot L_{12} \cdot i_1 i_2 \cdot \sin \alpha_{1-2} = 2\tau_P \cdot f_2 \cdot F_t. \quad (2.2)$$

Obviously energy is transferred from the primary to the secondary winding if  $f_2 \cdot F_t > 0$ . Hence energy transfer is possible not only at braking but also at accelerating and driving.

**Short-Stator Induction Motor [Yang (2004)]** The long-stator motor features high performance but it is relatively expensive. That is why the short-stator induction motor is investigated, too. At this motor the track-side long-stator is replaced by a simple aluminum/iron reaction plate, Fig. 17. At this cheap motor energy for driving must be delivered from the on-board power supply to the vehicle-side winding. By the traveling wave of this winding currents are induced into the reaction plate. Since no possibility exists to control phase displacement  $\alpha_{1-2}$  directly thrust force of this motor is lower than with long-stator motor.

Reaction plates are installed and short-stator operation is performed at approximately 30 m of the test track including the surrounding of the switch where installation of long-

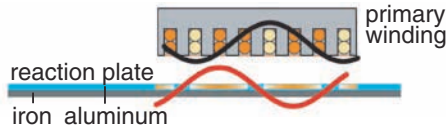


Figure 17: Short-stator induction motor

stator devices is difficult or impossible. When the linear motor is operated in the short-stator mode all energy for driving and supplying the hydraulic and electronic on-board units is delivered by an on-board battery.

## 2.2 Power Supply

**Power Supply of the Long-Stator [Yang and Grotstollen (2007)]** The long-stator is divided into separately fed sections consisting of series connected stator devices. When designing the power supply of sections two schemes have been compared. If the scheme of Transrapid and Maglev shall be applied a pair of power converters per vehicle and one switching unit per section are required as shown at Fig. 18a for independent operation of two vehicles. Alternatively a distributed supply has been considered at which each section is fed by a converter of its own, Fig. 18b. When doing the design for operation of three vehicles the distributed supply surprisingly proved to be less expensive because no switching units are required and a smaller number of shielded motor cables have to be connected.

At the test track all energy required on-board the vehicles is transferred via the linear motor and neither current rails nor overhead wires have been installed. Consequently the long-stator motor is an important part of the power supply which, as well as the feeding power converters, must have more rated power than required for driving only.

**Power Supply of the Vehicle [Pottharst (2006), Pottharst et al. (2004)]** On board of each vehicle a continuous and safe power supply is required for operation of the hydraulic system and all signal processing electronics. In addition energy is required during short-stator operation at which energy for driving must be applied to the vehicle-side motor part. Therefore energy is stored in a battery having great storage capability and ultra-capacitors which can handle high peaks of power occurring e. g. at braking, see Fig. 19.

To ensure continuous availability of on-board energy a dedicated energy management is under development which determines the reference of the power to be transferred by the linear motor and the distribution of this power to the battery and capacitors.

## 2.3 Control of Doubly-Fed Linear Induction Motor

**Communication for Distributed Control of Test Track [Pottharst and Grotstollen (2003), Pottharst (2006)]** Control of test track is distributed to a stationary control room, to vehicles moving on the track and to the stationary power converters

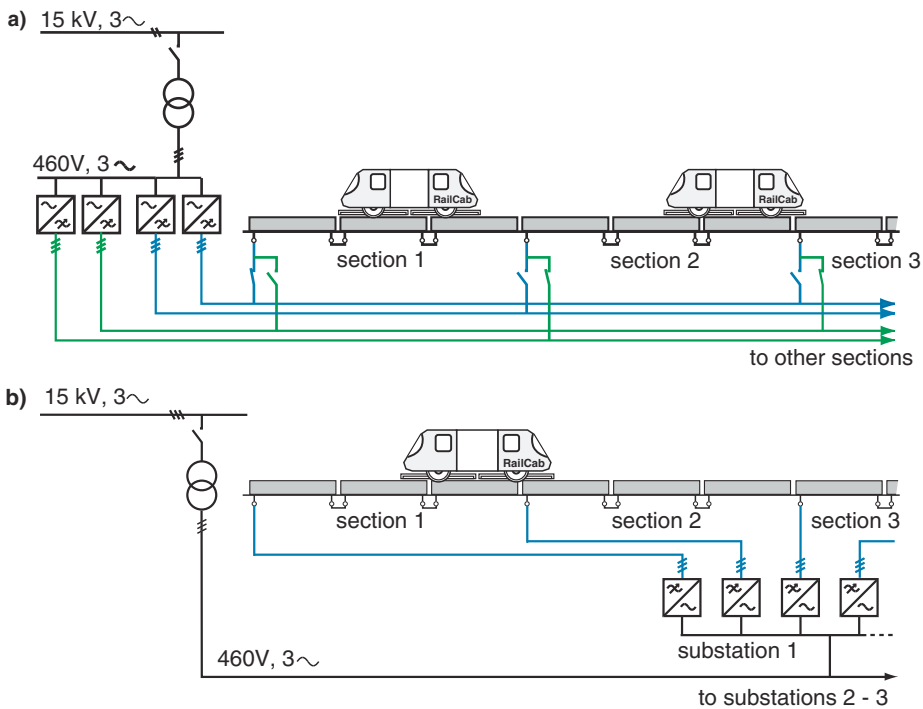


Figure 18: Power Supply of Long-Stator. a) Scheme of centralized power supply b) Scheme of distributed power supply

feeding the long-stator, see Fig. 20.

At the control room references for motion of shuttles are set and sent to vehicles via radio transmission. On board of each vehicle motion control is performed and references for currents are determined under consideration of requirements of energy transfer. References of primary current are sent to the track-side power converters via radio transmission and CAN bus. In case of more than one vehicle traveling on the same stator section this task is performed by one of them which was declared master. Additional radio communication takes place between vehicles moving on the track. Last not least direct access from the control room to the track-side converters is available.

### Distributed Drive Control [Pottharst and Grotstollen (2005), Pottharst (2006)]

Control of motion and energy transfer is performed on-board each vehicle individually, see Fig. 21. In accordance with commands for destination  $x_{V_d}$  and traveling speed  $v_{V_d}$  being received from the central stationary control room by radio transmission references for motion control are generated in the 'profile generator'. Depending on the situation reference for thrust force, speed or position is transferred to the block 'position and speed control' at which the actual values of position and speed are available for closed loop control. At the block 'operating point assignment' the references of primary and secondary

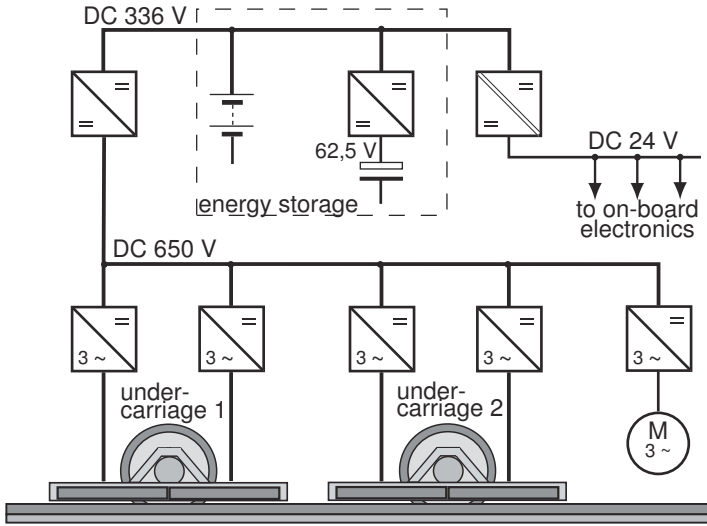


Figure 19: Structure of the on-board grid

current are determined under consideration of the thrust force to be generated and the power to be transmitted. The reference for power transmission is delivered by an 'energy module operating strategy' at which the requirements of the on-board power system is monitored and predicted continuously.

References of amplitude and frequency of primary current are sent to the power track-side power converters by radio transmission and CAN bus every 20 ms. With the hardware in use transmission causes a delay of 44 ms. To ensure exact control of angle  $\alpha_{1-2}$  and stability of control loops this delay is considered in the block 'synchronization' in which an estimate  $\hat{\epsilon}_1$  of the actual angle  $\epsilon_1$  of the primary current vector is determined. Closed loop control of currents is performed by the track-side and vehicle-side power converters as usual.

**Control of Thrust Force and Current Vectors [Henke and Grotstollen (2000), Grotstollen (2005), Henke (2003)]** For control of current vectors, a d/q-coordinate system was chosen which is oriented to the primary current. This choice proved to be suitable because the windings of all vehicles traveling on the same section of the long-stator have to interact with the traveling wave generated by this very current. Due to this choice of coordinate system the primary current vector  $\underline{i}_1$  has, by definition, no q-component in contrast with secondary current  $\underline{i}_2$

$$\underline{i}_1 = i_{1d} + 0, \quad \underline{i}_2 = i_{2d} + j i_{2q}. \tag{2.3}$$

According to (2.1) thrust force does not depend on the d-component of the secondary current which therefore is controlled to  $i_{2d} = 0$  meaning  $\underline{i}_2 = j i_{2q}$  and  $\alpha_{1-2} = \pm\pi/2$ . The stator current is common to all vehicles traveling on the same stator section and cannot be

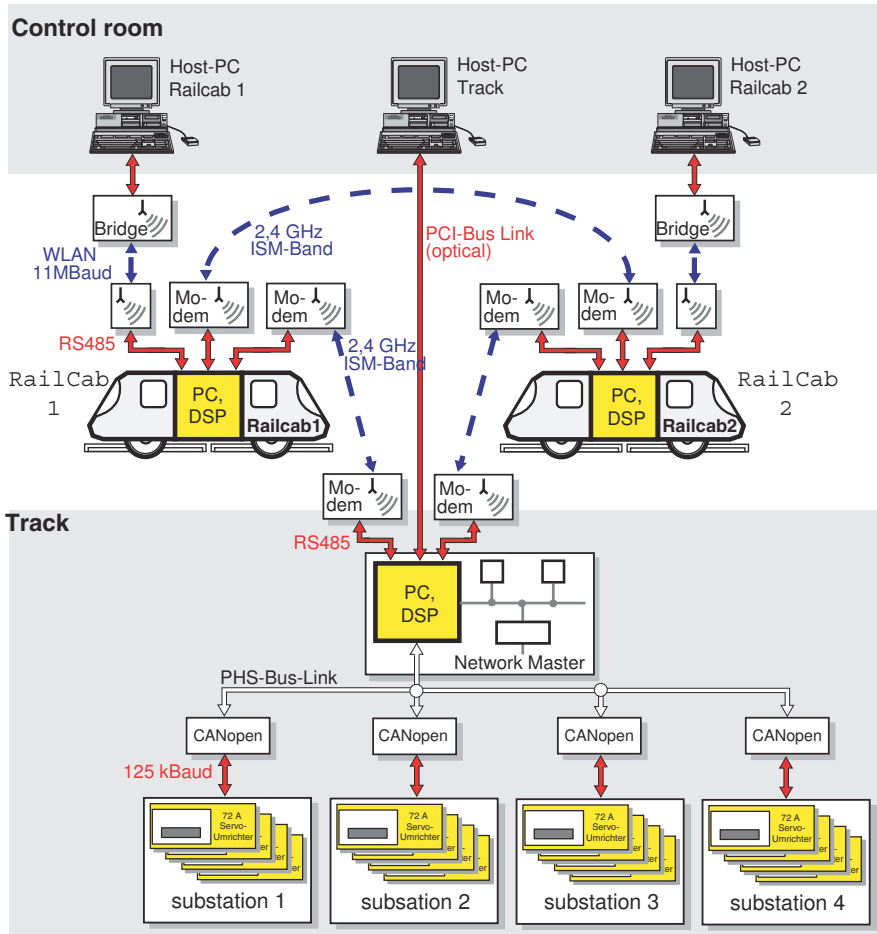


Figure 20: Distributed plant control

varied but with a great delay. Therefore thrust force, speed and position of each vehicle are controlled by means of its secondary current's q-component. For doing so position of the stator's traveling wave must be known with great accuracy because if the error  $\Delta\alpha_{1-2}$  exceeds  $\pm\pi/2$  thrust force changes its sign and control becomes instable. Due to orientation of the secondary current to the primary current the well-known frequency requirement

$$f_1 = f_2 + \frac{1}{2\tau_P} \cdot v_V \quad (2.4)$$

is fulfilled automatically. Satisfying this condition is a must at any three-phase machine because if traveling waves of primary and secondary winding are not synchronized, displacement  $\alpha_{1-2}$  would vary and thrust force  $F_t$  would not be but oscillating.



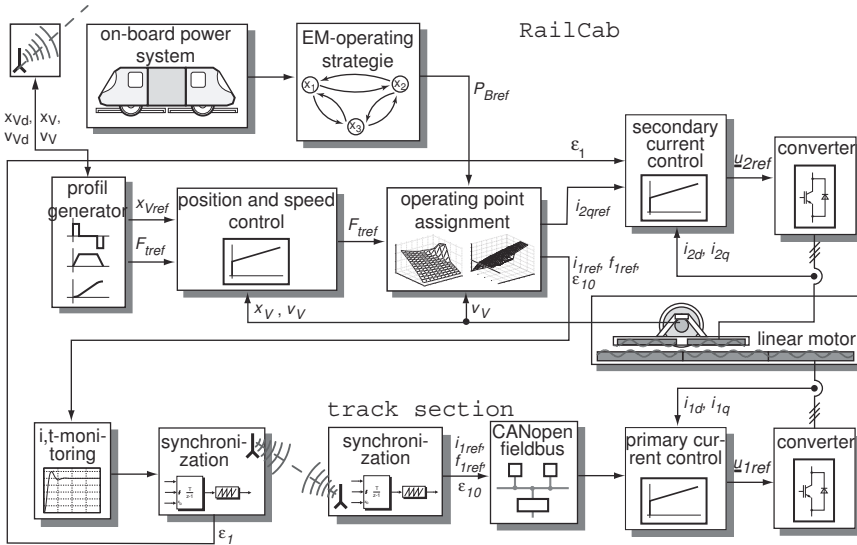


Figure 21: Scheme of control distributed to vehicle and track

Orientation of current vector  $\underline{i}_2$  perpendicular to the primary’s traveling wave cannot be performed without knowledge of the wave’s position which determines the d-axis. Hence the position of the primary’s traveling wave must be determined on board of the vehicle with an error much smaller than  $\Delta x_V = 50 \text{ mm}$  which refers to the critical angle of  $\Delta \alpha_{1-2} = \pi/2$ . When driving with maximum speed (10 m/s) the vehicle will cover this distance four times during one sampling time (20 ms) and almost nine times during the time required for data transmission (44 ms). Therefore synchronization of traveling waves, i.e. orientation of secondary field perpendicular to primary field, is a demanding task which is discussed in the following section.

## 2.4 Synchronization of Traveling Waves [Grotstollen (2006), Schneider et al. (2006)]

Control of wave displacement  $\alpha_{1-2}$  is performed by proper control of angle  $\epsilon_2$ . This angle represents the phase angle of current vector  $\underline{i}_2$  in the winding-oriented coordinate frame, and determines the position of the secondary’s traveling wave in this very frame, see Fig. 22.

Two possibilities for determination of the reference of  $\epsilon_2$  are investigated by which the position of the primary’s traveling wave is determined differently.

**Measurement of the Vehicle’s Position** In the first case the position of the primary’s traveling wave is calculated from the primary current’s phase  $\epsilon_1$  and the vehicle’s position  $x_V$  represented by angle  $\alpha_V$ . From Fig. 22 (left side) can be seen that the

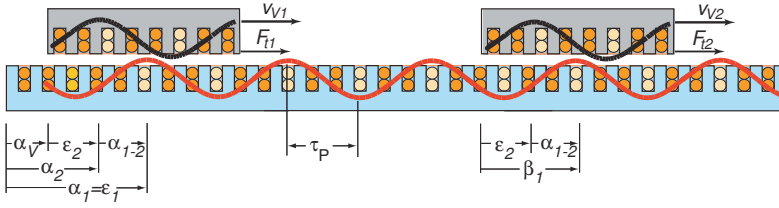


Figure 22: Determination of primary traveling wave

reference  $\epsilon_{2ref}$  of  $\epsilon_2$  can be calculated according to

$$\epsilon_{2ref} = \hat{\epsilon}_1 - \frac{\pi}{\tau_P} \cdot \hat{x}_V - \alpha_{1-2ref}. \quad (2.5)$$

In this equation  $\hat{\epsilon}_1$  is an estimated and  $\hat{x}_V$  a measured value while  $\alpha_{1-2ref} = \pm\pi/2$  and its sign is determined by the sign of the demanded thrust force. Considering that  $\epsilon_1$  and  $x_V$  can vary quickly it is important to determine  $\hat{\epsilon}_1$  and  $\hat{x}_V$  exactly for that instant for which  $\epsilon_{2ref}$  shall be calculated.

The position of the primary's traveling agrees with the angle  $\epsilon_1$  of the primary current for which an estimate is calculated in the block 'synchronization' of Fig. 20,

$$\hat{\epsilon}_1 = 2\pi \cdot f_{1ref} + \epsilon_{10}, \quad (2.6)$$

where  $f_{1ref}$  and  $\epsilon_{10}$  are updated every 20 ms under consideration of the 44 ms caused by radio transmission. By adding the vehicle's position, represented by the second term of (6), wave position is transformed to the vehicle-oriented frame. The vehicle's position  $x_V$  is measured by shaft encoders with high resolution of 0.31 mm which is required for determination of speed. Since 0.31 mm refers to a displacement of  $0.55^\circ$  accuracy of these sensors is very good but several effects can cause errors and must be considered. Only the most important phenomenon is mentioned here: Extremely great errors occur when the flange of a wheel strikes against the rail head.

To reduce the influence of the mentioned disturbances several sensors and measures have been implemented, see Fig. 23.

First and most important, the incremental measuring system is calibrated periodically by updating the initial value  $\epsilon_{10}$ . This is performed by evaluating the pulses delivered by a proximity switch which indicates markers installed at the front end of each stator device i.e. every 1200 mm. Periodical adjustment of position has proved to be important and must be performed with great reliability. Therefore a redundant measuring system has been introduced which is based on IR-sensors. These sensors are used to scan the surface of the stator segments and to detect the gap between neighbored segments. An advantage of these sensors is the reaction within microseconds but attention must be paid to the possibility of disturbance by stray light. Second, motion of all wheels is measured by shaft encoders and the average  $\hat{x}_V$  of measured values is used for control only if it is in accordance with the estimate  $\hat{\epsilon}_1$  calculated from (2.6).

Because of relatively simple realization this method of synchronization was implemented at first to accelerate commissioning.

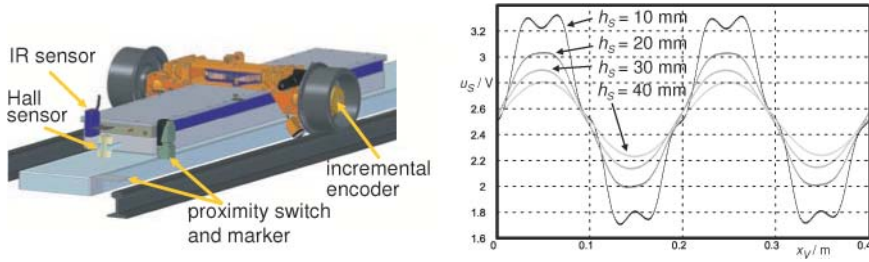


Figure 23: Measuring for synchronization. Left: Sensors at the vehicle. Right: Signal of Hall sensor (hS: height above stator)

**Measurement of the Primary Wave’s Position** To improve robustness and reliability of synchronization a second method is under development at which the position of the stator’s traveling wave is measured directly on board of the vehicle by means of Hall sensors, see Fig. 23. Since the position is measured in the vehicle-oriented coordinate frame (angle  $\beta_1$  in Fig. 22 calculation of  $\epsilon_{2ref}$  is as simple as  $\epsilon_{2ref} = \beta_1 - \alpha_1 - 2ref$ .

At the vehicle Hall sensors are mounted in a distance from the stator which is larger than the air gap. Hence they cannot be damaged or destroyed mechanically and errors caused harmonics, see Fig. 23 can be neglected.

## 2.5 Experimental Results [Pottharst and Grotstollen (2005), Pottharst (2006)]

At Fig. 24 measured step response of position and speed control is presented. During all the process shown reference of power to be transferred to the on-board supply system of the vehicle is 4 kW (1 kW for each secondary winding). As can be seen from the plots measured thrust force and speed match the references determined by the profile generator very well. In contrast the transferred power shows a considerable aberration from its reference which becomes extremely high after acceleration was terminated. Two reasons are most important: First, transferred power is not subject of a closed loop control. Second, after acceleration thrust force and related currents are decreased. Consequently voltages must be increased for transfer of constant power which is not possible when converters reach their limits.

Strong disturbances of thrust force can be observed after acceleration. They are caused by the rims of wheels touching the rails which happens because at the time of measuring steering mechanism was not in operation. The plots in Fig. 9 document prove the proper function of the synchronization mechanism and the drive control based on a distributed hardware structure. In spite of rapid changes of primary current and frequency a safe operation of the linear motor driven RailCab with energy transfer in the required dimension is possible.

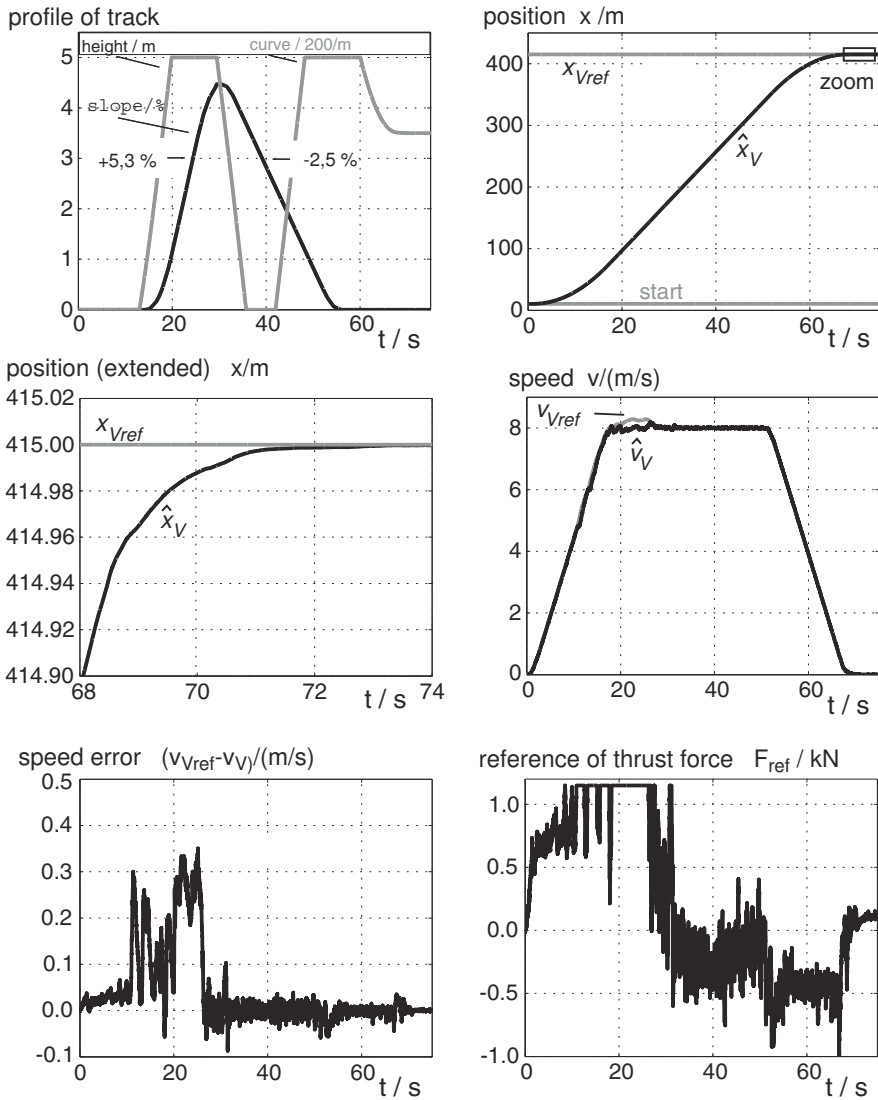


Figure 24: Step response of speed control with control of energy transfer

### 3 Development of Mechatronic Systems

Modern products like the Railcab are more and more complex. This complexity can only be contained by an appropriate structuring of the overall system into subsystems. Mechatronic systems are set up of components from different domains. Each domain structures its part of the overall system from a different standpoint. Thus, a control

engineer will structure the controller in a different manner than a mechanical design engineer the mechanical body. In this manner a high number of different system structures are created for one and the same overall system. For mechatronic systems, which take advantage of the harmonic interaction of the involved domains, this situation is dissatisfactory. They require a cross-domain, unique structure of the overall system according to mechatronic considerations, or in other words, a mechatronic aggregation, which forms the basic structure for all involved domains. This mechatronic aggregation also forms the macrostructure for the information processing of the mechatronic system.

### 3.1 Mechatronic aggregation - a function oriented approach

**Motional functions** As start point for the mechatronic aggregation the development of a function structure has proven its worth in many application. However, an appropriate understanding of the term “function” is a key enabler for deriving a useful mechatronic aggregate structure from the function structure. The Mechatronics Laboratory Paderborn (*MLAP*) has successfully used the following definition (see [Stolpe (2004), Lückel (2004), Hestermeyer (2006)]):

A function that describes the movement of a well-defined rigid body system shall be called a *motional function*. Motional functions form a subgroup of the purpose functions known from design methodology (Beitz and Pahl (1997)). Material-, energy- and signal-flow do not need to be considered; they are an implicit result of the function oriented aggregation (see Figure 25).

Mechatronic systems are characterized by realizing motional functions. In order to do so, they have to fulfill mechanical, sensorial, actuator and software functions (Naumann (2000)) with components from different domains. These four *elementary functions* are typical for mechatronic systems and can be found in one or the other way in all mechatronic systems. The principal idea behind the function structure for the mechatronic aggregation consists in limiting the function structure (in a first step, see below) to motional functions, without mentioning the elementary functions explicitly. This prevents domain specific structures within the overall function structure.

In order to visualize the effect of partitioning a system by means of motional functions, it is helpful to take a look at the rigid-body model of the overall system. The motional function structure divides the overall mechatronic system into partial mass systems with their own controlled motions. There is a 1:1 relationship between a motional function and a partial mass system.

In certain situations it might become necessary to further detail motional functions by *internal motional functions* that all refer to one and the same partial mass system. It is e.g. possible to group the lift, row and pitch body motion of an automobile on the one hand and the yaw, lateral and longitudinal motion on the other hand. Both groups describe motions of the same partial mass system.

**Mechatronic aggregates for motional functions** The aggregation of the mechatronic system can be derived from the described function structure by allocating mecha-

tronic aggregates to each function in the structure. These aggregates have to include components for the realization of the elementary functions given above. Depending on the position of the function in the hierarchy, the aggregates have a different internal set-up. It is possible to distinguish between two basic aggregate types:

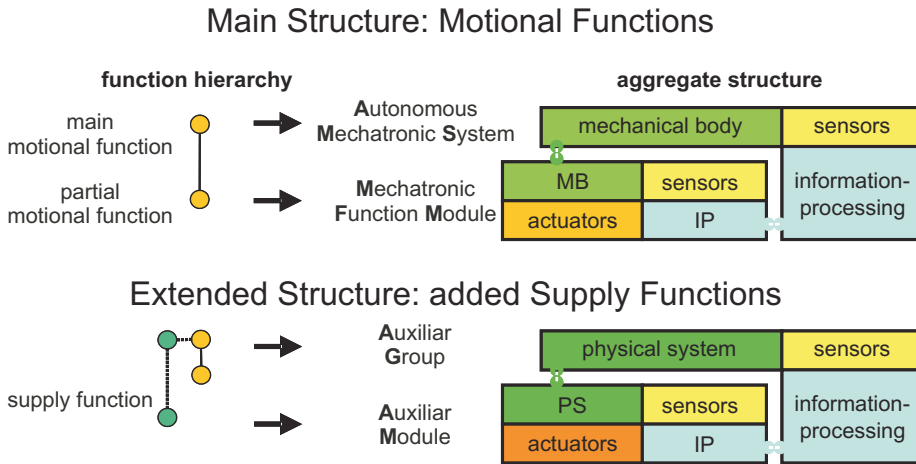


Figure 25: Function structure and aggregation.

- Mechatronic Function Modules - MFM* MFM form the base elements of mechatronic systems and include components for the realization of all four elementary functions (Lückel (1992)). This results in the well-known set-up of a mechatronic system (see Figure 25), as presented e.g. by Honekamp et al. (1994). The motion of the mechanical body is measured by sensors and controlled by actuators to fulfill the motional function. The loop is closed by information processing elements. Actuators can themselves again be MFMs. A typical example of an MFM is an electronically controlled hydraulic cylinder.
- Autonomous Mechatronic Systems - AMS* Honekamp et al. (1997) call the aggregate for the realization of the upper-most motional function in the function structure an AMS. It contains all aggregates with mechanical coupling and can interact autonomously within its environment. It therefore requires its own energy supply (Naumann (2000)). The AMS uses MFMs as actuators, so that there are no dedicated separate actuators for the AMS alone. The Railcab is a typical example of an AMS.

Internal motional functions that relate to the same partial mass system can not be realized as separate full-fledged aggregates. However, sensors and information processing of a MFM or the AMS can be grouped according to the internal motion functions. Lückel et al. (2001) call these groups *Mechatronic Function Groups - MFG*.

**Supply functions and their aggregates** For the complete structuring of complex mechatronic systems it is necessary to include auxiliary purpose functions apart from the motional functions in the function structure. These auxiliary purpose functions need to be “headline” functions similar to the motional functions, which imply elementary functions without naming them in the function structure. One important group of auxiliary functions is the group of the *supply functions*. Supply functions ensure energy supply or energy distribution within a technical system. It is not their task to control the energy flow in a way that a certain motion is realized - this is the task of actuators. They only ensure that the actuators can work. Ideally, there is no dynamic influence of the supply functions on the motion functions.

In the aggregation of the mechatronic system the supply functions are realized by Auxiliary Modules and Auxiliary Groups (see also Figure 25):

- *Auxiliary Modules - AM* The AM generalizes the MFM. It uses digital information processing to control the behavior of a physical system. The internal component structure of the AM corresponds to the structure of the MFM except that a general physical system takes the place of the mechanical structure. The borders between this physical system and the actuators are fluent. This is however not relevant for the realization.
- *Auxiliary Groups - AG* Different AMs - potentially from different domains like pneumatics, hydraulics, etc. - are combined to AGs. AGs are “administrative units”: The different AMs within an AG do not (necessarily) have functional physical connections. They could however be mounted to a common rack. A combination of a hydraulic and pneumatic supply unit is a typical example for an AG. Information processing elements coordinate the different AMs and offer a common interface of the AG to external aggregates. In order to fulfill this coordination task, the AG may have its own sensors.

### 3.2 Aggregation and information processing

All aggregate types for the mechatronic aggregation described above feature a module for information processing. As the overall mechatronic system is composed of these aggregate types, the overall information processing of the mechatronic system is composed of the individual information processing modules of the different aggregates. Thus, each aggregate in the Railcab structure shown in Figure 27 also corresponds to an information processing module.

In the approach used for the Railcab, the communication between the aggregates mirrors the function respective aggregate structure and the mechanical/physical linkage. Thus, an aggregate communicates only to its parent aggregate, but not to other aggregates on the same level or other aggregates somewhere else in the hierarchy. Figure 27 visualizes how communication lines are chosen parallel to the coupling lines. The difference in the coupling and the functional connection between MFM/MFG/AMS and the AM/AG are reflected in different communication types (see section 5).

The module and communication structure thus obtained form a macrostructure for the information processing of a mechatronic system, which is applicable for the control scheme, the diagnostic monitoring and the system operation (see section 5).

### 3.3 Mechatronic aggregation of the Railcab

**Motional function structure** According to the above considerations, the starting point for the determination of the Railcab aggregation is the function structure with the main motional function at its top. In the case of the Railcab, the main motional function can directly be derived from the purpose function of the Railcab: Transporting passengers and goods as comfortably and cheaply as possible from one location to another. As passengers and goods are located in the vehicle body during transportation, this purpose function can be stated more precisely: Control the motion of the vehicle body in all six spatial degrees of freedom.

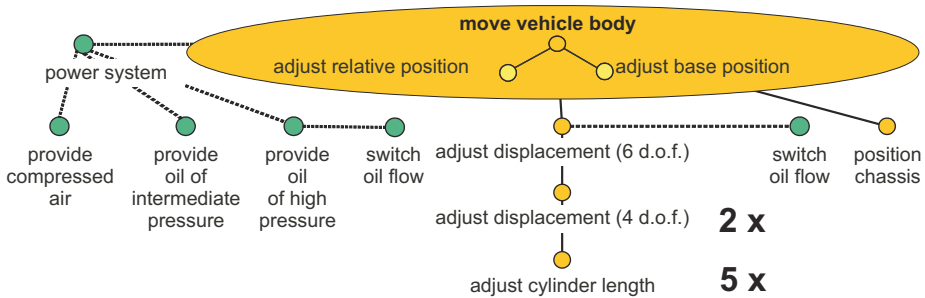


Figure 26: Railcab function structure (only hydraulic & pneumatic supply).

Looking at different types of railway vehicles, the principle realization of this motion function is quite similar:

A suspension system allows for small relative movement between vehicle body and chassis (formed by the bogeys) in all six degrees of freedom. Passive suspension systems control this relative movement by a clever choice of spring-damper-characteristics; fully active suspension systems provide an additional active control mechanism. The leeway for the essential, long movement along the track is given by the chassis. This motional function also includes the guidance of the chassis and the vehicle within the track.

According to these considerations the motion of the vehicle body is separated into a base motion and a relative motion. As base and relative motions relate to one and the same mass system, they correspond to internal motional functions (see Figure 26), to be realized by MFG.

In the next step partial motional functions need to be identified, which can later on be realized as MFM. The determination of these partial motional functions requires a further elaboration of the main motional function as part of the mechatronic composition (Töpper (2002) identifies the mechatronic composition as important design step during which the principle solution for a design task is determined by means of model analysis using simplified solution elements.). For the Railcab, this is shown in Figure 26. The base position is adjusted by the active chassis. The relative motion of the body is made possible passively by the air springs. In order to provide a means for the active control of the relative motion, the active tilt technology actively creates a relative displacement between the upper air spring joint and the vehicle body in all six degrees of freedom. The



base position adjustment and the creation of this relative displacement form the partial motional functions to the main motional function.

Both partial motional functions can be further detailed. Figure 26 indicates this for the relative displacement in six degrees of freedom. It is split into two displacements in 4 d.o.f. (displacement within a plane and longitudinal displacement of this plane) and finally into one-dimensional displacements, which can be realized by hydraulic cylinders.

**Supply functions** The motional functions are complemented by supply functions, which depend on the concrete solution for the actuators. The Railcab requires three supply functions:

- The basic energy source is electrical power from the batteries, which are charged and discharged depending on the driving situation. However, this function is currently not yet integrated into the overall function oriented mechatronic concept of the Railcab and thus not shown in Figure 26.
- In order to allow for the motion between vehicle body and chassis, air springs are used. These require a compressed air supply.
- Most actuators in the Railcab use hydraulic energy, thus e.g. the active tilt system. Due to the hydraulic setup, two different pressure levels are necessary: one main pressure level and one intermediate level at above half the main pressure level.

In order to keep the system structure simple, the supply-functions are bundled under a supply-function “power system”.

**Aggregate structure and macrostructure of the information processing** The aggregate structure can be derived from the function structure. The internal motional functions are realized by MFG, the partial motional functions by MFM. Supply functions are fulfilled by AM and AG. The aggregate structure to Figure 26 is shown in Figure 27.

As described above, this aggregate structure is mirrored in the macrostructure of the overall system’s information processing. Even though the Railcab currently uses centralized hardware that controls all modules, this macrostructure has been consequently used for development and setup of the overall information processing.

### 3.4 Design Steps for IP-modules within the macrostructure

The previous paragraph has shown how a characteristic macrostructure for the information processing of the overall mechatronic system can be derived from function and aggregate structure. The IP-modules of the macrostructure can be realized either on centralized or (preferably) on distributed hardware. Deriving the macrostructure from the function structure is the starting point for the development of the mechatronic system and its information processing. These structures are determined “top-down”.

**MFM & AMS** The elaboration of MFM and AMS usually starts with the lowest MFM in the aggregate structure. It is therefore done “bottom-up”. The elaboration of the information processing follows this process, as it is closely linked to the elaboration of the aggregate (see Figure 28).

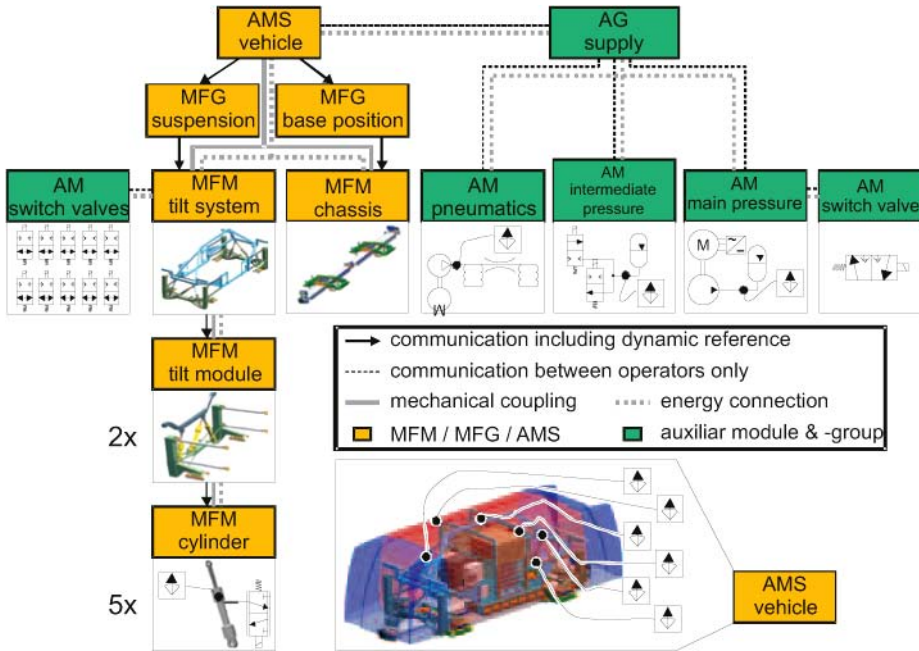


Figure 27: Railcab aggregate structure (only hydraulic & pneumatic supply).

1. The design of the information processing starts with the lowest IP-module within the macrostructure. Each IP-module is tuned and parameterized for the corresponding *separate* MFM. That is to say that the MFM is not mounted to the assembly it is usually located in. This is made possible by the mechanical independence of the MFM, whose task it is to control the motion of a *dedicated, one-to-one* mass system. This independence also allows for a commissioning of the MFM separate from the overall system and further in-depth examinations without stressing the mechanical structure of other modules.
2. In a next step, several MFM are combined to form a higher-order aggregate. Due to changes in the MFM control plants and interactions between the controllers of the different MFM, the dynamical behavior of the MFM usually changes compared to step 1. Therefore it becomes necessary to check the function of each MFM respective its IP-module also when mounted into the higher-order aggregate. This is the case as well for model-based design as for commissioning of the physical system.
3. During the third step, the IP-module of the higher-order aggregate is elaborated. The subordinate MFMs including their information processing are part of the control plant. The controller structure and the parameters of the subordinate information processing are kept fixed during this step.
4. The last step comprises the development of information processing for the overall

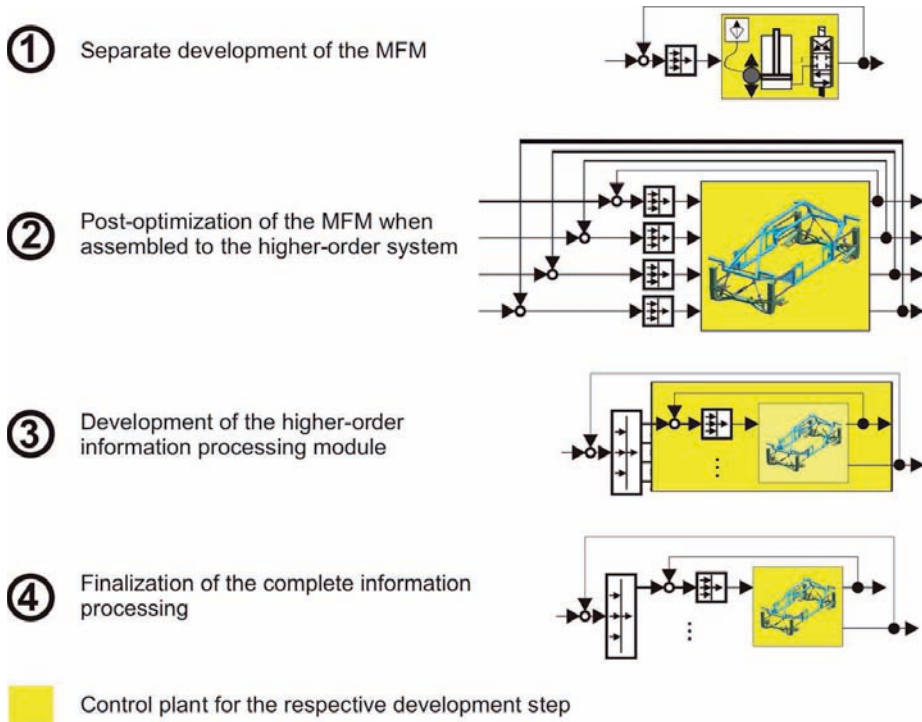


Figure 28: Design steps for IP-modules within the macrostructure.

system. This includes post-optimization of parameters as well as the introduction of additional functions like e.g. disturbance compensation with MFM-external signals, which are only possible by the interaction of different aggregates. Post-optimization includes the parameters of the subordinate aggregates. This ensures, that the dynamical coupling between different subordinate aggregates caused by their integration in a higher-order aggregate is considered in the optimal parametrization (see also [Lückel et al. (2001)]).

In case the overall system of step 4 is itself again part of a high-order aggregate, the development continues with step 2.

**Supply** If the supply modules are designed in a way, that they are approximately decoupled from the consumers in their dynamic behavior, their IP-modules can be elaborated independently from MFM and AMS. This is the case for the Railcab supply modules. The elaboration within the hierarchy of the supply aggregates is performed bottom-up, similar to the process of MFM and AMS. As supply modules of higher-level MFM are often necessary for the commissioning of low-level MFM, it is a good idea to elaborate and start them up in an early stage of the overall development, regardless of their position within the hierarchy of AMS and MFM.

### 3.5 Verification of dynamic coupling between MFM

**The macrostructure as generalized cascade** The aggregation of mechatronic systems defines as described in the above chapters a characteristic macrostructure for the information processing of the mechatronic system. Looking from a control engineers point of view, this macrostructure forms a characteristic control configuration with a horizontal and vertical decomposition. The dominant characteristic of this configuration is the vertical, hierarchical decomposition, the control cascade. Lückel et al. (2001) call this configuration therefore a generalized cascade. Different from conventional cascade controllers as described e.g. by Skogestad and Postlethwaite (1996), a controller higher up in the hierarchy sets reference values not only for one, but for several different subordinate controllers. This results not only in a hierarchical, but also in a decentralized decomposition.

The hierarchical controller design and step-by-step start of operation, beginning with the lowest control loops in the hierarchy, requires, that lower-level MFM need to work independently from other MFM of the same layer, but in parallel - without any coordination superior controller (see Figure 28, step 2). During this phase, the control configuration is of a decentralized nature, so that controllers in the generalized cascade have in parts to fulfill requirements common to decentralized control structures. This requirement on independent operation of the MFM also supports the idea of modularization: When a MFM A fails, there should be no instability of a MFM B. This should also not be the case if the communication fails.

The commissioning and robust operation of a generalized cascade therefore requires certain independence between the different modules. Depending on the mechanical coupling however, this independence is only given partly or not at all, so that their might be interaction between different modules, that cause instable behavior even though each MFM is stable for itself. In other words: It is possible that an MFM works fine if operating alone. But as soon as a second MFM is turned on, the system becomes unstable.

Within a central controller, this instability can be avoided by e.g. introducing decoupling control elements. From a point of view of modularization, this is however not possible. Thus, configurations with a strong coupling between modules must be avoided; the function and aggregate structure must then be changed towards a more decoupled setup.

**The relative gain array** In order to evaluate, whether a plant is controllable in a decentralized way or it is likely that significant difficulties will be encountered, the linear control theory has developed a technique that quantifies the decentralized controllability. This technique can be employed to check the aggregate structure for its feasibility.

Bristol (1966) introduced the *relative gain array* (RGA) as measure for the interaction between different plant inputs and outputs. It is the objective of the RGA to determine, how a transfer path from one input  $u_j$  to an output  $y_i$  changes, when other transfer paths are controlled. Bristol distinguishes between two extreme cases:

- All other paths are left uncontrolled, that is  $u_l = 0 \mid l \neq j$ . The transfer function of the path  $j \rightarrow i$  can be computed by  $\left( \frac{\partial y_i}{\partial u_j} \right)_{u_l=0, l \neq j}$ .

- All other paths are “ideally” controlled. The set point for these control loops is 0. Therefore  $y_l = 0 \mid l \neq i$ . The transfer function of the path  $j \rightarrow i$  can be computed by  $\left(\frac{\partial y_i}{\partial u_j}\right)_{y_l=0, l \neq i}$ .

The ratio of both cases of one transfer path  $j \rightarrow i$  yield a measure for the change of the transfer behavior due to control of other transfer paths. The matrix  $\underline{\underline{\mathbf{A}}}(s)$  of the ratios of all transfer paths is called RGA:

$$(\underline{\underline{\mathbf{A}}}(s))_{ij} = \left( \left(\frac{\partial y_i}{\partial u_j}\right)_{u_l=0, l \neq j} / \left(\frac{\partial y_i}{\partial u_j}\right)_{y_l=0, l \neq i} \right)$$

According to Skogestad and Postlethwaite (1996), for a given transfer matrix  $\underline{\underline{\mathbf{G}}}(s)$ ,  $\underline{\underline{\mathbf{A}}}(s)$  can be evaluated to equation (3.1):

$$\underline{\underline{\mathbf{A}}}(s) = \underline{\underline{\mathbf{G}}}(s) \times (\underline{\underline{\mathbf{G}}}(s)^{-1})^T \tag{3.1}$$

$\times$  denotes element-wise multiplication, also called Schur-product.

If the model description in the transfer matrix is arranged in a way, that (assuming single-input-single-output-MFM) input and output of each MFM can be found at the same position in the input vector  $\underline{u}$  and the output vector  $\underline{y}$  of the overall plant, so that therefore  $u_i$  and  $y_i$  belong to the same aggregate, then in the ideal case the RGA should be  $(\underline{\underline{\mathbf{A}}})_{ii}(s) = 1 \forall s$ . This means, that the transfer behavior of an aggregate does not change when control loops of other aggregates are closed or opened. If the values of the RGA differ significantly from 1 especially in the region of the cross-over frequency, there is a strong link between the respective aggregate and other modules: Controlling aggregate A changes the transfer function amplitude of module B. Transfer paths with  $(\underline{\underline{\mathbf{A}}})_{ii}(0) < 0$  are extremely hard to control: In these cases the start of other aggregates results in a change of sign of the transfer function. Configurations including such RGA values should absolutely be avoided.

When developing the function and aggregate structure of the mechatronic system, it should be ensured that  $(\underline{\underline{\mathbf{A}}})_{ii}(s) \approx 1$  in the frequency range of interest. This is especially important for supply aggregates. If this is given, no significant interactions between controllers of different MFM need to be expected. The RGA thus gives the chance to check, whether the chosen aggregate structure is sensible also from a control engineering point of view - and that independently from a specific controller design.

**Relative gain array for the MFM cylinder** The application of the RGA shall be demonstrated for the tilt system of the Railcab. As shown in Figure 27, the tilt system uses 10 hydraulic cylinders for generating an active displacement between vehicle body and the upper couple point of the air springs.

In a first step, a linearized model of the suspension tilt system has been developed, that includes electrical amplifiers, a hydraulic model containing capacities, inductivities and valves, a rigid body model containing 7 masses and sensor models. From this model, a  $10 \times 10$  transfer matrix  $\underline{\underline{\mathbf{G}}}(f)$  as function of frequency  $f$  has been computed. The control voltages to the current amplifiers for the hydraulic valves that control the cylinders

constitute the input to this matrix. 10 cylinder length measurements are the output. The transfer matrix  $\underline{\underline{G}}(f)$  thus forms a linear model for the behavior of the MFM cylinder after integration into the tilt system.

Using  $\underline{\underline{G}}(f)$  (3.1) yields the RGA  $\underline{\underline{\Lambda}}(f)$ . Figure 29 depicts the values of  $\underline{\underline{\Lambda}}(f)$  in the complex plane for frequencies between 0.1Hz and 20Hz and six of the 10 cylinders, which move the body in vertical direction.

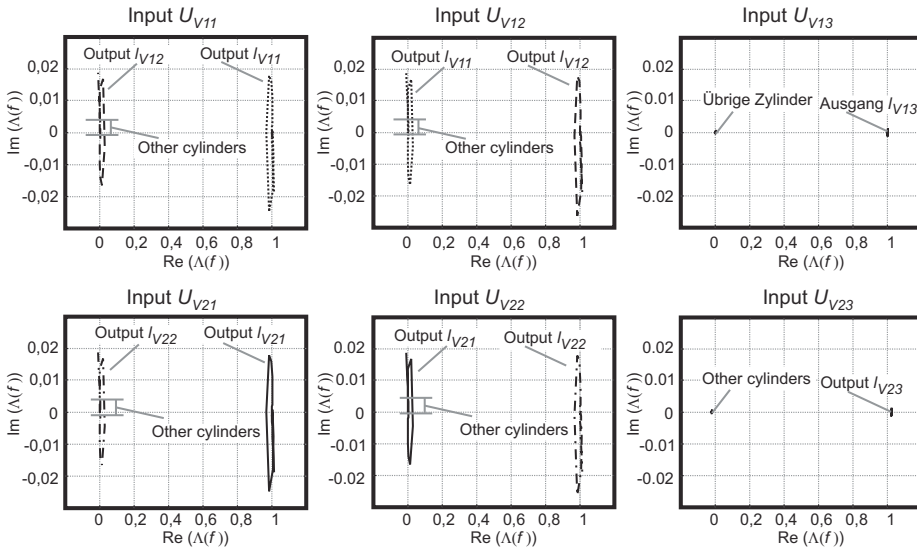


Figure 29: Relative Gain Array for the hydraulic cylinders.

The RGA shows a nearly complete independence between the cylinder modules. The matrix elements  $(\underline{\underline{\Lambda}})_{ii}(f)$  for input and output of one cylinder module are virtually constant 1 in the frequency range till 20Hz, matrix elements between inputs and outputs of different cylinder modules are virtually 0. There is therefore no destabilizing interaction between the controllers. The chosen modularization is convenient also from the point of view of control engineering.

## 4 Model of HiL-Testbed, Suspension/Tilt System

### 4.1 Introduction

The aim in designing the testbed is to analyse the entire configuration of the suspension/tilt system by measurement techniques under real-time conditions and to demonstrate the novel active suspension/tilt module in a realistic environment. The theoretical models can be identified by real components; in this manner a model-based design of the control becomes possible in the first place. The hierarchical controller structures designed can be implemented on the testbed and analysed, tested, and validated by measurement techniques under real-time conditions. Moreover, fundamental research on actuator and

sensor systems for the NBP test shuttle is done [Liu-Henke (2005)]. In the following, we will present the concept of the suspension/tilt module that the build-up of the testbed is based upon. Then we will deal with the model-based design of the testbed - from the modelling to the model-based controller design to the realisation of the testbed.

## 4.2 Concept of the Suspension/Tilt Module

An essential requirement of modern traffic engineering is ride comfort. One uses the acceleration of the carriage body to assess the ride comfort because acceleration is felt as a disturbance by the passenger. That is why carriage-body accelerations should be as small as possible. This can be achieved by isolating the carriage body from the disturbances the carriage is subject to by a soft suspension in both horizontal and vertical directions.

Another objective is utmost ride safety; it can be obtained by using dampers. A system with dampers is subject to high damper forces in the upper frequency range. Thus, interception of the damper force generates additional acceleration of the carriage body, again diminishing ride comfort.

The result is a conflict of goals in passive suspension systems that are equipped with hard spring/damper units. This conflict can be solved by an active suspension where the spring and damping forces can be actively adjusted at will [Lückel (2000)].

Figure 30 displays the concept to be presented. In contrast to conventional suspension systems, the concept presented is concerned with eliminating all passive dampers in the secondary suspension and connecting the carriage body to the undercarriage merely via the airsprings. Thus the desired soft suspension can be obtained, which brings about the optimal decoupling of the carriage body from the undercarriage in the degrees of freedom in question. While the airspring isolates the vibrations in the upper frequency range, the desired damping in the lower frequency range is obtained by an active displacement of the base of the airspring by means of the upper plate [Liu-Henke et al. (2000)]. Thus the transmission of disturbances to the carriage body stemming from an uneven track alignment is near zero. The result is a more than satisfactory ride comfort both in the vertical and horizontal directions. The information required to control the base displacement is measured by the appropriate sensors and processed in a hierarchically structured multi-variable control system.

The active tilting of the carriage body can be realized with the same actuator system through a feedforward of the cornering acceleration. The actuator system is laid out for both these tasks and has enough bandwidth (5-10 Hz) to ensure the actuators being sufficiently dynamical to operate the tilting system while at the same time fulfilling their suspension task.

The core of the actuator system comprises the following mechanical components: upper and under plate resp. the airsprings, the actuators, and the sensors. While the actuators A and B do the lifting and tilting, the third actuator, C, mainly provides for lateral motion.

Thus every suspension/tilt module has to ensure three directions of motion: vertical, lateral, and tilting motions. The shuttle (consisting of two undercarriages) allows all controlled rotational and translational motions in lateral and vertical direction. The translational motion in the longitudinal direction is realized by means of the linear drive

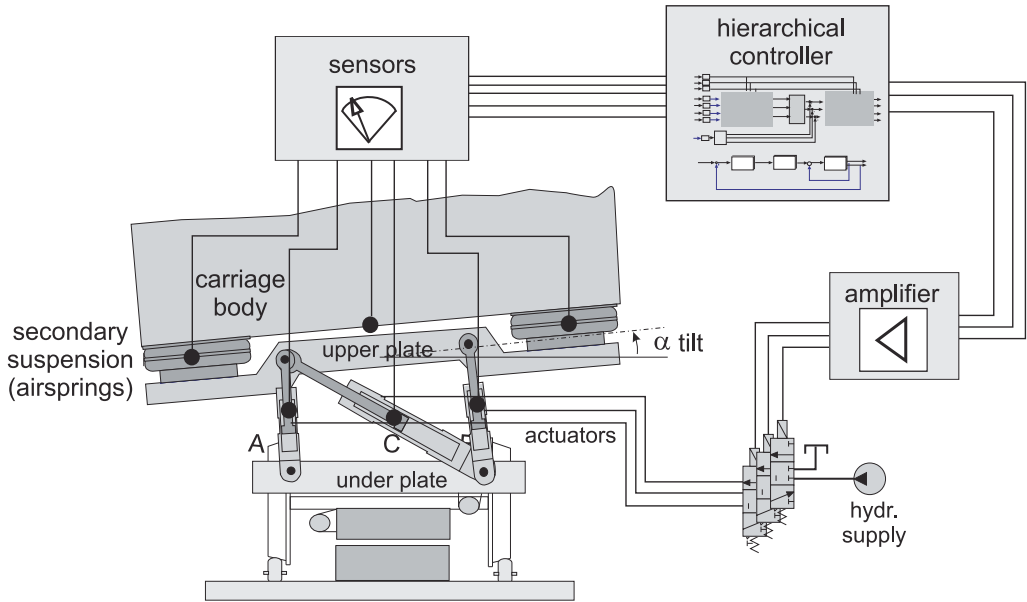


Figure 30: Concept of the suspension/tilt system

[Henke (2003)].

### 4.3 Modelling of the Suspension/tilt Module

In order to analyze the behaviour of a dynamical system and then to design a multi-variable control, the physical and the mathematical substitute models must be drawn up according to functional principles in view of the demands. The model has to represent the kinematic, static and dynamical behaviour of the system in question [Liu-Henke (2005)].

#### *Kinematic Functions*

The kinematics has to perform analysis on the kinematic conditions in the suspension/tilt system and to derive symbolically the forward kinematics and inverse kinematics of the overall system.

On the one hand, the symbolical-mathematical understanding of the system kinematics is a vital element of the entire control concept and will be embedded into the control components of the carriage-body control (chapter 4.4: Overall Control Structure) that need the kinematics for the base-displacement of the airspring. On the other hand, symbolic derivation of the forward and inverse kinematics makes a realization of the control structures in a real-time application easier. Moreover the working area of the system can be determined by an analysis of the kinematics.

The kinematic behaviour of the system is determined by the degrees of freedom and the geometry of the suspension/tilt module. For this purpose, one has to choose appropri-



ate coordinate systems. Figure 31 displays them: the ICS (Initial Coordinate System) on the upper edge of the track, the BCS (Body Coordinate System) in every rigid body, the ACSs (Attachment Point Coordinate Systems) in the coupling points between interlinked elements, and the coordinates of every cylinder piston  $x_{cylinder}$ .

The forward and inverse kinematics of the entire system are calculated by means of a coordinate transformation and will subsequently be integrated into the global control loops.

**Dynamic Functions**

The dynamic functions describe the dynamical behaviour under the effect of forces. Thus we can make a mathematical statement on the system behaviour using the principles of physics. Figure 31 illustrates the physical substitute model of the suspension/tilt module, in this case the dynamics of a multi-body system with corresponding actuators, sensors, and information processing.

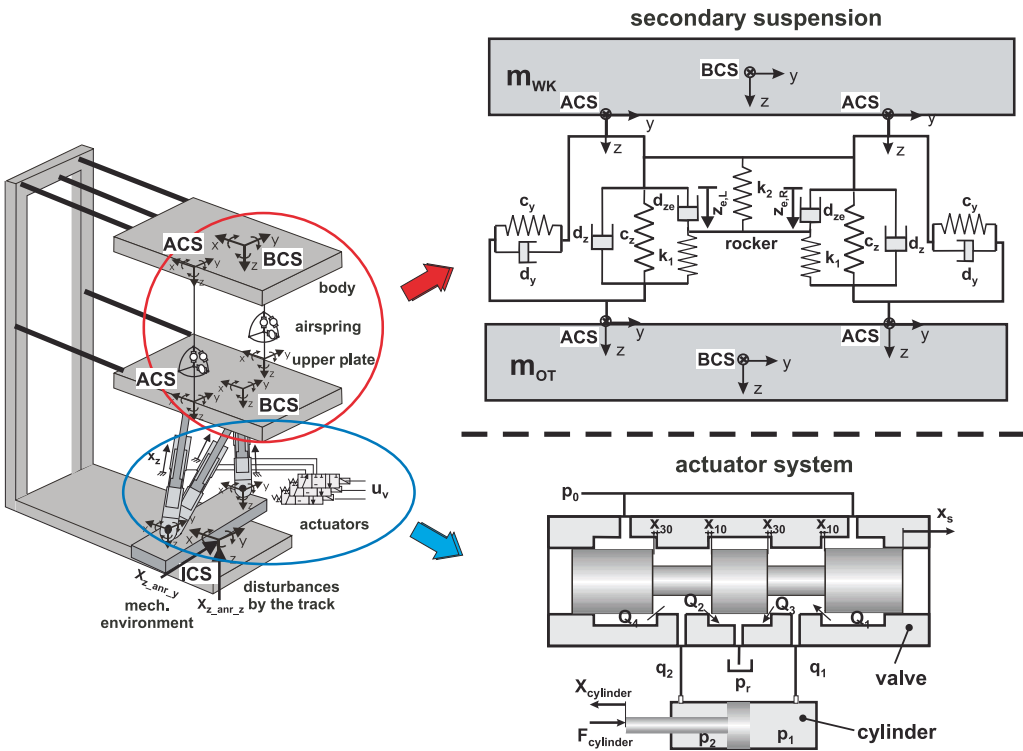


Figure 31: Physical substitute model

**Dynamics of the Secondary Suspension** The carriage body of the suspension/tilt module is modelled as a rigid body with six degrees of freedom (DOFs). The trailing link installed in longitudinal direction (see Fig. 31: mechanical guidance) that is meant

to reduce the DOFs in the system is regarded as a mechanical linking element. Thus the secondary suspension has three DOFs ( $y_{wk}$ ,  $z_{wk}$ ,  $\alpha_{wk}$ ). For a mechanical environment we have the upper edge of the rail that the carriage body is connected to via the mechanical linking element (mechanical guidance). The airspring force affects the carriage body and the upper plate according the kinematics.

Due to the construction the exchange of air between bellows and the supplementary air container is very small in lateral direction; thus in lateral direction the pneumatic stiffness can be neglected. The airspring force results mainly from the deformation of the bellows [Liu-Henke and Ise (2002)]. Thus spring-damper behaviour of airsprings can be modelled with the spring constant  $c_y$  and the damper constant  $d_y$  that represents the dynamical behaviour in lateral direction.

The vehicle dynamics in lifting and tilting directions are considerably affected by the exchange of air bellows and the supplementary air container. The airspring force results not only from the mechanical suspension ( $c_z$ ) and the damping ( $d_z$ ) exerted by the bellows in vertical direction but also from the air compression that is modelled by a pneumatic stiffness of the air enclosed in the bellows ( $k_1$ ). Due to the flow resistances in the throttle, a damping capacity determined by the throttle diameter emerges along with the spring force. Thus a damper ( $d_{ze}$ ) which is interconnected with spring ( $k_1$ ) in series is employed. We take into account that the two airsprings share a the supplementary air container. There is only one the supplementary air container, yet two throttles are used for the airspring system. The effect of the massless spring of the supplementary air container ( $k_2$ ) is distributed to the two springs by a rocker.

The airspring forces are dependent of the relative motion forces (displacement and velocity of displacement) between at the carriage body and the upper plate:

$${}^{ACS}\vec{F}_{LFL} = \underline{\underline{c}} \left( {}^{ACS}\Delta\vec{r}_{LFL} - {}^{ACS}\vec{r}_{0,L} \right) + \underline{\underline{d}} {}^{ACS}\underline{v}_{LFL} + \vec{F}_{0,L} \quad (4.1)$$

$${}^{ACS}\vec{F}_{LFR} = \underline{\underline{c}} \left( {}^{ACS}\Delta\vec{r}_{LFR} - {}^{ACS}\vec{r}_{0,R} \right) + \underline{\underline{d}} {}^{ACS}\underline{v}_{LFR} + \vec{F}_{0,R} \quad (4.2)$$

**Dynamics of the Actuator System** The suspension/tilt actuator system comprises above all the upper plate and the three actuators. The upper plate is modelled in the shape of a rigid body with six DOFs. As is the case with the carriage body, the upper plate is connected to the environment via a longitudinal control arm as a mechanical linking element; thus the DOFs of the upper plate are reduced to the lateral motion  $y_{up}$ , the lifting motion  $z_{up}$ , and the tilting  $\alpha_{up}$ .

As displayed in Figure 31, the actuator system consists of three differential hydraulic cylinders with three servo valves. The cylinder chambers are operated separately by one valve each. The valve transforms the electrical signal dynamically via the motion of the gate valve into a pressure-oil flow and thus regulates the motion of the cylinder supplying the actuator force to the system.

**Dynamics of the Valve** The dynamic behaviour of the individual servo valve describes the way of the gate-valve displacement  $x_s$  is dependent of the electrical input voltage  $u_V$ ;

it can be described approximately by a system of 2nd order:

$$\ddot{x}_S + 2\frac{d_V}{T_V}\dot{x}_S + \frac{1}{T_V^2}x_S = \frac{1}{T_V^2}u_V \quad (4.3)$$

The statical properties of the servo valve, bringing about high non-linearities in the component, are described by the flow-rate characteristics. Here we have a four-edge-gate valve with symmetrical negative underlap of the control edges ( $x_{10}$ ,  $x_{30}$ ).

If the Bernoulli equation is employed, the flow equations describe to which extent the volume flow rate ( $Q_i$ ) of each control edge depends on the position of the gate valve ( $x_s$ ):

$$Q_1 = K_V \cdot \left( \frac{-x_{10} + x_S + |-x_{10} + x_S|}{2} \cdot \sqrt{|p_0 - p_1|} \cdot \text{sgn}(p_0 - p_1) \right) \quad (4.4)$$

$$Q_2 = K_V \cdot \left( \frac{-x_{10} + x_S + |-x_{10} + x_S|}{2} \cdot \sqrt{|p_2 - p_r|} \cdot \text{sgn}(p_2 - p_r) \right) \quad (4.5)$$

$$Q_3 = K_V \cdot \left( \frac{-x_{30} + x_S - |-x_{30} - x_S|}{2} \cdot \sqrt{|p_1 - p_r|} \cdot \text{sgn}(p_1 - p_r) \right) \quad (4.6)$$

$$Q_4 = K_V \cdot \left( \frac{-x_{30} + x_S - |-x_{30} - x_S|}{2} \cdot \sqrt{|p_0 - p_2|} \cdot \text{sgn}(p_0 - p_2) \right) \quad (4.7)$$

$$K_V = Q_{nenn} \cdot \sqrt{\frac{2}{\Delta p_{nenn}}} \quad , \text{-valve gain} \quad (4.8)$$

The balance yields the volume flows being transmitted to the cylinder:

$$q_1 = Q_1 - Q_3 \quad (4.9)$$

$$q_2 = Q_4 - Q_2 \quad (4.10)$$

**Dynamics of the Cylinder** The cylinder converts the difference in pressure stemming from the oil flow into and out of the valve into an actuator force in the two cylinder chambers and transmits the force to the upper plate.

The dynamics of the individual cylinder depends on the hydraulic capacities  $c_{h1}$ ,  $c_{h2}$  and on oil leakage rates. The mass of the piston of the cylinder is not taken into account here because the resulting natural frequencies are considerably above the frequency range in question.

The law of continuity yields the mathematical state equations with the states chamber pressure  $p_1$ ,  $p_2$ . As system inputs, there are the volume flows  $q_1$ ,  $q_2$  and the velocity of the piston  $\dot{x}_{cylinder}$ :

$$\text{chamber 1: } c_{h1} \cdot \dot{p}_1 = q_1 - a_1 \cdot \dot{x}_{cylinder} \quad (4.11)$$

$$\text{chamber 2: } c_{h2} \cdot \dot{p}_2 = -q_2 + a_2 \cdot \dot{x}_{cylinder} \quad (4.12)$$

The result is the actuator force of every single cylinder with regard to the friction force  $F_r$ :

$$F_{cylinder} = p_1 \cdot a_1 - p_2 \cdot a_2 - F_r \quad (4.13)$$

### Dynamics of the Sensor System and of the Digital Information Processing

When designing a controller, one has to make sure that the overall dynamics takes into account the dynamical behaviour of the sensors employed and the dead times resulting from the controller's being realized digitally.

For the position sensor an inductive sensor is chosen to trace the displacement of the cylinders and airsprings. The position of each gate valve is measured by a displacement pick-up element integrated into the servo valve [Liu-Henke et al. (2000)]. The sensors have dynamics of their own, an effect to be reckoned with at the development stage and to be modelled by means of a PT2 (a second-order delay element) low-pass filter.

In the realization of a digital controller, the sampling and the processing time necessary to compute the control algorithm can be described by a dead time in the millisecond range. For this purpose, a Padé approximation is implemented to emulate these digital effects.

#### 4.4 Model-based Design of the Cascading Control Structure

Based upon the identified mechatronic model of the vehicle [Liu-Henke and Ise (2002)] the control structures can be laid out in a model-based manner so that the desired system behaviour can be achieved. For this purpose we apply a method developed at the MLAP, the generalised cascade principle [Lückel et al. (2001)], on the design of the multivariable control of the suspension/tilt module.

Figure 32 displays the cascaded control structure and its interfaces. The superordinated, central controller group is employed for controlling the carriage-body dynamics while the local controller improves the dynamics of every single actuator and processes it decentrally. The interface of the hierarchically organized control structure is made up of the three reference positions of the cylinders  $x_{z\_ref\_A}$ ,  $x_{z\_ref\_B}$ ,  $x_{z\_ref\_C}$ .

**Overall Control Structure** It was shown above that the dynamics of the secondary suspension is a multi- variable problem where the three carriage-body motions are strongly interlinked due to their physical nature. The problem lies in the design of the multi- variable control for an intermeshed system. In order to achieve more clarity in the system behaviour and do an easy controller design with as few control variables as possible, we aim at separating the entire system into loosely connected subsystems [Liu-Henke et al. (2002)].

The basic idea in designing the global controller is to separate the carriage-body motions into the respective modal degrees of freedom by means of a modal decoupling. This may considerably facilitate the design of the multivariable control for now it is possible to consider every modal motion separately and to design one optimal controller individually for every motion. On this basis a state controller in combination with the skyhook damper is employed for the control of every single motion [Liu-Henke (2005)].

Therefore the superordinated global controller consists basically of three blocks: **decoupling, tuning forces, and coupling**.

The **decoupling** has to convert the measured displacements of the cylinders and airsprings into quasi-modal variables of the carriage body (vertical and lateral displacements, tilting angle) via the direct kinematics of the system.

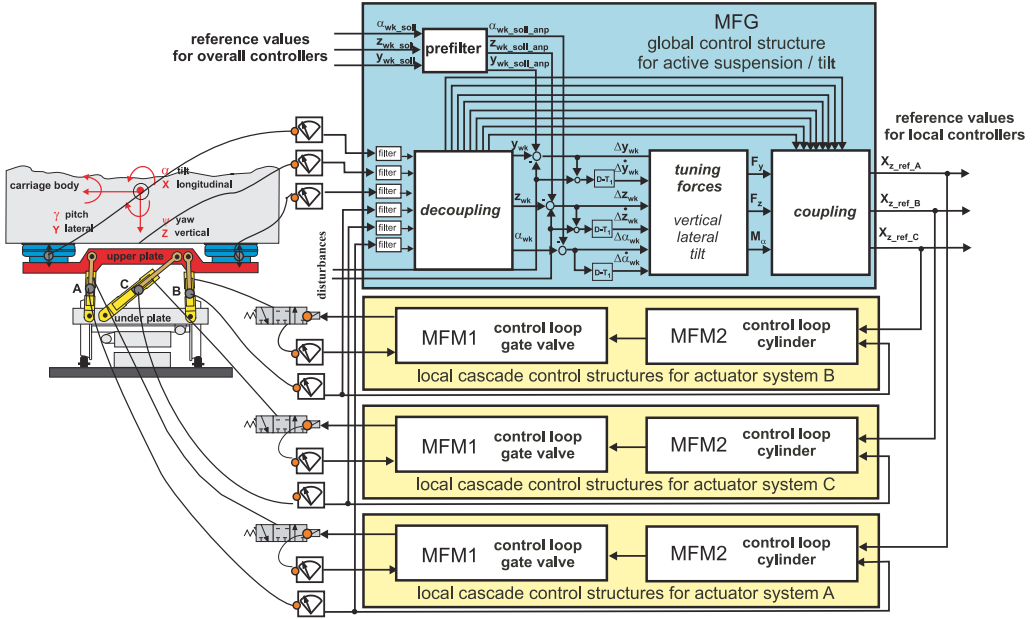


Figure 32: Cascaded multi-variable control structure of the suspension/tilt module

The core process of the carriage-body control is the **tuning force**. From the reference position and the current carriage-body position  ${}^{ICS}\underline{r}_{wk,ist} = (y_{wk} \ z_{wk} \ \alpha_{wk})^{-1}$  yielded by the decoupling we can determine the control difference resp. the spring-deflection velocities  ${}^{ICS}\dot{\underline{r}}_{wk,ist} = (\Delta\dot{y}_{wk} \ \Delta\dot{z}_{wk} \ \Delta\dot{\alpha}_{wk})^{-1}$  derived from it. On this basis the desired forces  $\underline{F}_{soll} = (F_y \ F_z \ M_\alpha)^{-1}$  for the active struts with predetermined vertical, lateral, and rolling-motion springs resp. dampers are generated in the *tuning force* block using the skyhook principle.

$$\underbrace{\begin{pmatrix} F_y \\ F_z \\ M_\alpha \end{pmatrix}}_{\underline{F}_{soll}} = \underbrace{\begin{pmatrix} c_y & 0 & 0 \\ 0 & c_z & 0 \\ 0 & 0 & c_\alpha \end{pmatrix}}_{\underline{C}} \cdot \begin{pmatrix} \Delta y_{wk} \\ \Delta z_{wk} \\ \Delta \alpha_{wk} \end{pmatrix} + \underbrace{\begin{pmatrix} d_y & 0 & 0 \\ 0 & d_z & 0 \\ 0 & 0 & d_\alpha \end{pmatrix}}_{\underline{D}} \cdot \begin{pmatrix} \Delta \dot{y}_{wk} \\ \Delta \dot{z}_{wk} \\ \Delta \dot{\alpha}_{wk} \end{pmatrix} + \underbrace{\begin{pmatrix} F_{y0} \\ F_{z0} \\ M_{\alpha 0} \end{pmatrix}}_{\underline{F}_0} \quad (4.14)$$

The desired forces  $\underline{F}_{soll}$  depend on the desired spring rate  $\underline{C}$  and the desired damper constant of the skyhook damper  $\underline{D}$ . The elements of the matrices  $\underline{C}$  and  $\underline{D}$  represent the spring stiffnesses and the dampings that have to be tuned for three controlled motions:

- desired suspension rate and desired damping rate in lateral direction  $c_y, d_y$
- desired suspension rate and desired damping rate in vertical direction  $c_z, d_z$
- desired suspension rate and desired damping rate in the direction of the rolling motion  $c_\alpha, d_\alpha$ .

The tuning of controller parameters is a matter of model-based synthesis that is based on the identified model of the suspension/tilt module and executed off line. Therefore the six controller parameters are determined in the entire nonlinear model with the optimisation tool MOPO, an add-on to CAMEL-View.

Then the new reference values ( $x_{z\_ref\_A}$ ,  $x_{z\_ref\_B}$ ,  $x_{z\_ref\_C}$ ) for the subordinated control loops (at the local struts) are computed from these modal forces as well as the forces resulting from the cornering acceleration. This is done in the **decoupling** block by means of the inverse system kinematics.

**Local Control Structure** As local controls there are the cascade control structures with two loops. The inner control loop is made up of the position control of the gate valve, based on the measurement of the gate-valve displacements. This control loop has the smallest time constant and provides sufficiently high dynamics for the valve.

The outer control loop is made up of the position control of the cylinder. The displacement of the cylinder piston is measured and further processed as a state feedback in the controller. The position controller generates the reference signal for the subordinated valve control loop.

#### 4.5 Hardware-in-the-Loop Testbed

The theoretical studies detailed above served as the basis for mounting the testbed to examine the suspension/tilt module under real-time conditions. The submodels are successively replaced in the entire mechatronic system by corresponding real components; this is the origin of the hardware-in-the-loop testbed.

Figure 33 shows the realized testbed of the suspension/tilt module. It makes up a half-shuttle with a suspension/tilt module on a scale of 1 : 2.5 [Liu-Henke et al. (2002)]. The carriage body has three degrees of freedom: vertical and lateral as well as tilting motions. The excitation cylinders in the test infrastructure emulate the vibrations resulting from the primary part of the undercarriage and from unevennesses in the track bed, as well as from lateral disturbances on the carriage body.

The positions of the cylinder, gate valve, and the displacements of the airspring are measured. The measured signals are transmitted to a digital signal processor by A/D converters. There the control algorithms are processed and the signals transmitted to the actuating valves of the real subsystem via D/A converters.

The testbed was successively put into operation, level by level, with the most virulent (fastest) level, MFM1, of the controlled valve being implemented first and then the level MFM2 of the controlled cylinder. When the motions of the three actuator systems were harmonized to an optimum, the MFG “active suspension/tilt” was implemented. Thus the entire system was successfully put into operation.

On this testbed the component parts and the entire configuration of the system can be analyzed and validated under real-time conditions; the control structures designed according to the above procedure were further optimized under these conditions.

Figure 34 contrasts the simulation and measurement data both in the time range (left-hand side) and in the frequency range (right-hand side).

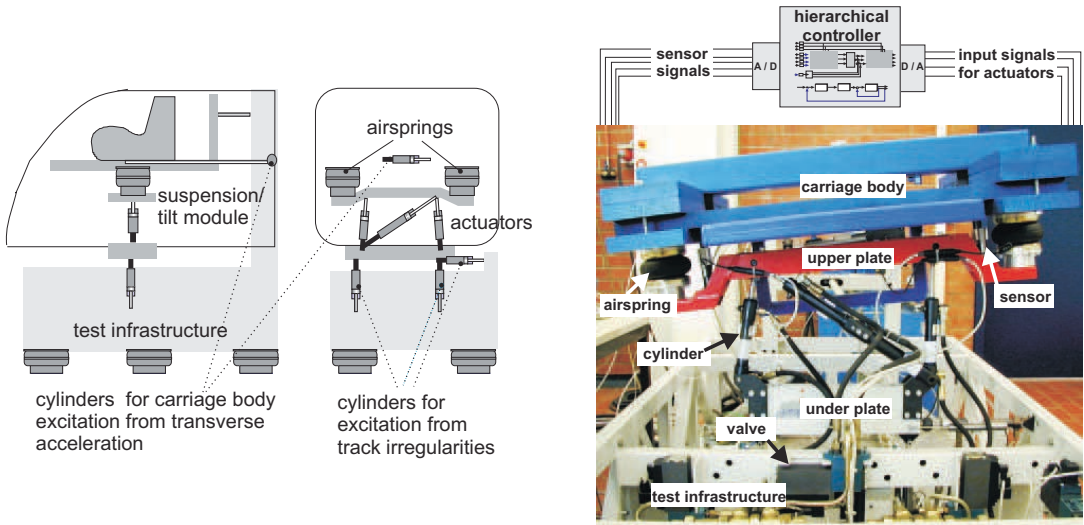


Figure 33: Realization of the testbed

The results in the time range represent the carriage-body motions of the active and the passive systems in vertical and lateral directions. Measurements at the testbed revealed the vibrational behaviour of the active system to boast a far better damping than that of the passive system and the system dynamics, compared to that of the passive system, to be considerably improved by the hierarchical controller structure. The active system boasts an excellent guiding and disturbance reaction of the controlled system.

The areas under the amplitude spectrum are a measure of the accelerations that the passenger is subject to. The larger the area, the worse the ride comfort provided by the system. In order to display the potential inherent in the active system for improving ride comfort, we take the passive system realised on the testbed and a calculated conventional railway vehicle for a reference.

Figure 34 contrasts all acceleration amplitude responses of these three vehicles as regards vertical and lateral direction. These amplitude responses clearly reveal the control to function perfectly over the entire frequency range (up to 20 Hz). The irregularities in the track can no longer be transmitted to the carriage body and a considerable decoupling of the disturbance is achieved. The active suspension provides ride comfort that is three times better than that of the passive system and more than five times that of a conventional vehicle.

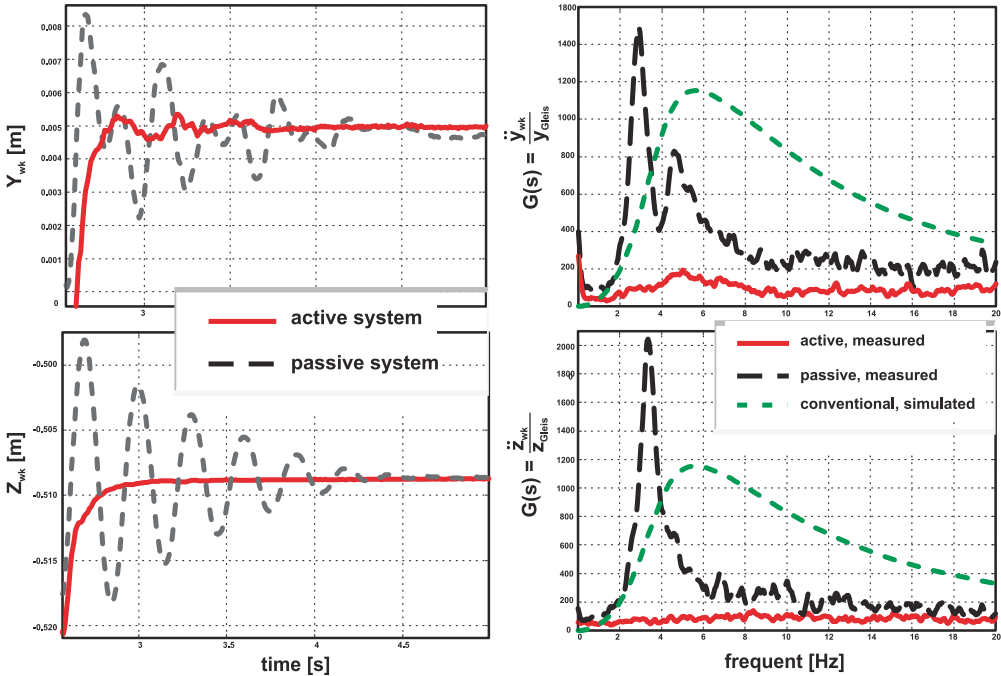


Figure 34: Results in the time and frequency ranges

## 5 Micro-Structure of the Information Processing

### 5.1 The Operator-Controller-Module

The macrostructure of the information processing described in section 3 divides the information processing of the overall mechatronic system into independent, separate modules. Such a module has to fulfill quite different tasks like the dynamic control of the associated aggregate, the sequencing of start/stop operation or the diagnostic monitoring of the aggregate. An appropriate microstructure of the macro modules can simplify the design task of the information processing by offering a clear template. The Railcab uses for microstructure so called *Operator-Controller-Modules* (OCM).

Figure 35 shows the OCM, which was also presented e.g. by Hestermeyer et al. (2004). Its set-up is oriented towards the kind of effect on the technical system:

1. On the lowest level of the OCM there is the controller. This innermost loop processes measurements and produces control signal. It directly affects the plant. It can therefore be called “motor” loop. Software processing on this level works in a quasi-continuous mode, i.e., measured values are read in, processed, and output continuously and under hard real-time conditions. The controller can be made up of several controllers, with the possibility to switch between them. The switch is done in one step; fade-over mechanisms and the like are again integrated in a separate controller element.



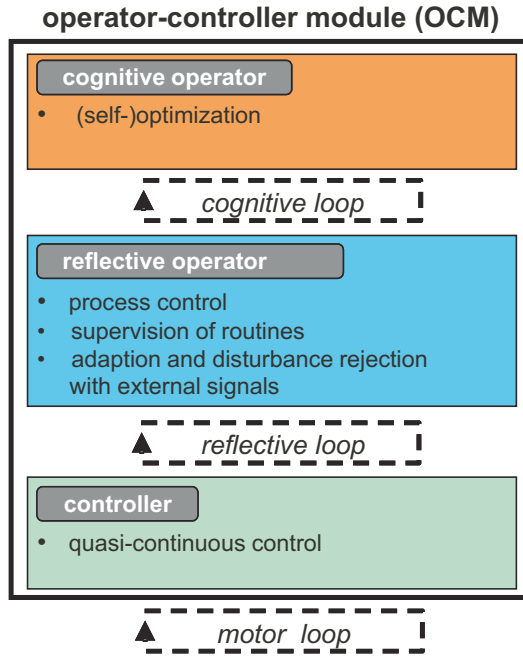


Figure 35: Operator-Controller-Module.

2. The controller is complemented by a reflective operator in which monitoring and controlling routines are executed. The reflective operator does not access the actuators of the system directly but modifies the controller. In certain circumstances it can also switch between different controller configurations. The reflective operator operates very much in an event-oriented manner. On this level, though, there are also other quasi-continuous functions, such as continuous adaptation algorithms or diagnostic monitoring functions. Because it is tightly linked to the controller, the reflective operator has also to operate under hard real-time constraints. The reflective operator, connecting element to the cognitive level of the OCM, provides an interface between those elements that are apt for real-time operation or operate in soft real time and the controller. It filters the incoming signals and inputs them to the subordinated levels.
3. The topmost level of the OCM is occupied by the cognitive operator. On this level the system can gather information on itself and its environment by applying various methods such as learning, use of knowledge-based systems, model-based optimization, and the like; it can employ them for improving its own behavior. Moreover, one can think of other cognitive functions. The present paper will confine itself to local self-optimization. This optimizing information processing can roughly be divided into model-based and behavior-based self-optimization. Model-based optimization allows an optimization that is predictive and decoupled in time from

the real system. The behavior-based optimization comprises functions for planning and evaluating the current objectives (cf. Oberschelp et al. (2002)). While both the controller and the reflective operator are subject to hard real-time constraints, the cognitive operator can also operate asynchronously to the real time. Of course it has to respond within a certain time limit; otherwise, due to altered environmental conditions, self-optimization would not find utilizable results. So the cognitive operator is subject to soft real time.

To sum things up, one can detect two distinct levels of separation: on the one hand, information processing is divided into two loops affecting the system, one directly and the other only indirectly. This division reflects the one that distinguishes between operator and controller. On the other hand, one can distinguish between hard and soft real time constraints. This classification sets the cognitive operator apart from the reflective one and the controller. As the level of separation is chosen according to the task we propose three separate elements.

## 5.2 The OCM for the active suspension of the Railcab

The operator-controller module of the MFG “suspension” in Figure 25 is a good example of the OCM microstructure: As described above, the information processing of the MFG has to implement a sequence controller and diagnostic monitoring functionality as part of the reflective operator (realized by a state machine with an error coder). In addition, the dynamic control of the body motion must be implemented. According to Goodall and Kortüm (1990), this task can be split into the following four parts:

1. Dynamics / Damping: Active suspension systems can improve the dynamic response of the vehicle body to track excitations by e.g. using measurements of the body acceleration to increase the damping of the body response (so called “sky-hook-damping”).
2. Tilt: Active tilt in curves reduces the effective stationary lateral acceleration of the passengers in curves. As this lateral acceleration is the limiting factor for the curve speed, active tilting enables faster curve speeds.
3. Lateral centering: An active system can center the vehicle body during curve transit in lateral direction, so that the end buffers are not reached (so called “Hold-off”). A further, for train operators even more important aspect is the possibility to use the reduced railway loading gage in curves to enlarge the wagons.
4. Level control: In addition, an active system is able to control the body level independently from the vehicle load also in vertical direction. This facilitates entering the vehicle as the entrance level can be adjusted to the platform level.

In order to determine, how these tasks can be realized in the OCM, it is important to keep the given suspension set-up in mind, which is shown in Figure 36, left.

The track excitation causes the vehicle body (in blue) to move. It is a noise input to the suspension system. The actual movement of the vehicle body in all six degrees of freedom is picked up by 6 accelerometers. For actuation, the active tilt system allows for a displacement between the air spring couple points and the vehicle body. In addition,

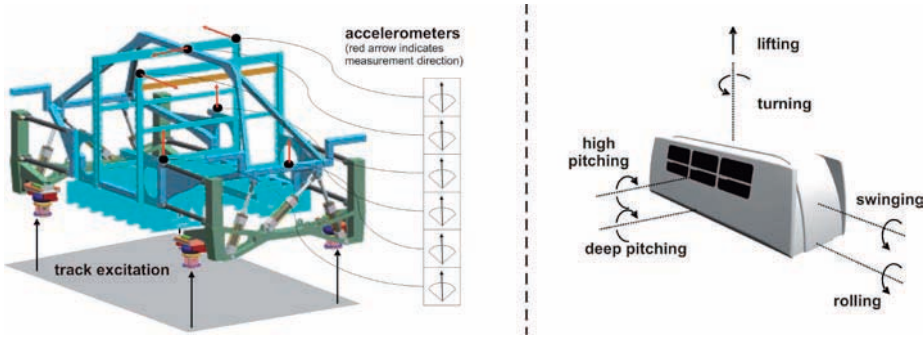


Figure 36: Set-up of the Railcab suspension (left) and natural motions (right).

the position and velocity of the Railcab are quite well-known, as they are necessary for running the linear motors. They are proprietary signals to the MFM “chassis”.

Restricting the suspension sensors to accelerometers has the advantage of low system costs and simple system set-up, but partly reduces functionality. Of the four tasks for the dynamic control named above, only three can be realized with these sensors: active damping, tilt and lateral centering. Level control is not possible. In addition, tilt and lateral centering can only be realized using non-suspension signals. These signals from the chassis have a different quality than those of the MFG “suspension” - they are only available in later stages of the development and have different timing properties.

Due to the different signal properties and the different dynamic link to the suspension plant behavior, the information processing for active damping, tilt and lateral centering are realized in different parts of the OCM: The Controller uses the acceleration-signals of the MFG “suspension” for closed-loop control, the reflective operator completes the control by using the non-suspension signals “vehicle speed”  $v$  and “position”  $s$  for a disturbance feed forward in open loop.

The overall OCM for the MFG “suspension” is shown in Figure 37.

The necessary tilt angle  $\alpha$  is derived from (5.1)

$$\alpha = \frac{v^2 \rho(s)}{g} \tag{5.1}$$

$g$  denotes the acceleration of gravity. The curvature  $\rho(s)$  is stored as look-up table dependent on the vehicle position  $s$ . As axis of tilt a line at the likely position of the passengers’ stomachs has been chosen, which has been found to be a good means against traveling sickness. The success of this rather straight-forward compensation can be seen in Figure 38, right. The results from a test run in a curve (active damping turned off) with a velocity of 5 m/s show that - if the position of the vehicle is known well enough - a feed-forward compensation of the lateral acceleration is well possible. As there are no dynamic interactions between tilt and damping control, feed-forward compensation is preferable to feed-back in this case.

As the necessary lateral displacement can be computed similar to (5.1), it shall not be further discussed here.

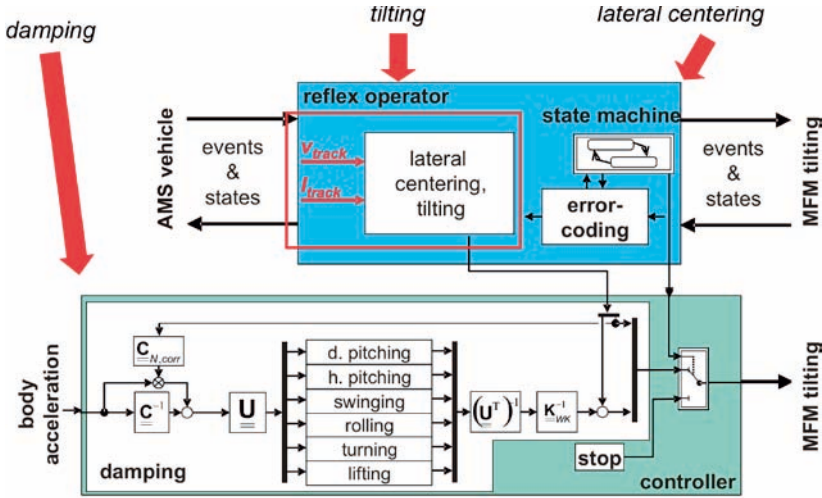


Figure 37: Operator controller module of the MFG “suspension”.

The dynamic damping control in the controller computes a reference displacement for the MFM tilting that is necessary to generate the required damping forces at the air springs, using a sky-hook-damping approach. This computation is done in two steps: In a first step, the necessary damping forces and torques are computed from the body acceleration signals, in the second step, the necessary displacements are derived, using the stiffness matrix  $\underline{\underline{K}}_{WK}$ .

The computation of the damping forces and torques is done in modal coordinates in order to simplify the tuning of the parameters: As vehicle chassis and body are connected only by air springs with negligible damping properties, it is possible to derive a linear model of motion for the generalized coordinates  $\underline{x}_B$  of the body in the form of (5.2)

$$\underline{\underline{\Theta}} \ddot{\underline{x}}_B(t) + \underline{\underline{K}} \underline{x}_B(t) = \underline{F}_{active}(t) + \underline{F}_{passive}(t) \tag{5.2}$$

with stiffness tensor  $\underline{\underline{K}}$ , inertia tensor  $\underline{\underline{\Theta}}$ , the excitation force  $\underline{F}_{passive}(t)$  caused by movement of the chassis and the active force  $\underline{F}_{active}(t)$  generated by the MFM “tilting”. Solving the generalized eigenvalue problem of the homogenous part of the differential equation yields six real eigenvalues and real eigenvectors, which decouple the equations and can be used as modal coordinates. Due to the fact, that the eigenvectors are real, they can easily be interpreted as natural translation and rotation motions of the body after external excitation. These natural motions of the vehicle are shown in Figure 36, right.

In order to prove the functionality of the active suspension, several test runs were performed on a defined part of the test track with active damping turned on and off. The active guidance module was deliberately turned off during these runs. Due to the stochastic nature of the excitation, the measurement results were evaluated by computing the standard deviation of the acceleration measurements, similar to Streiter et al. (2001).

As the active suspension is effective up to approximately 10Hz, higher frequency content was eliminated from the measurements using 3rd-order Butterworth filtering. Figure 38 compares the results from the test runs with and without active damping and visualizes the area, in which the standard deviations of the different test runs lay.

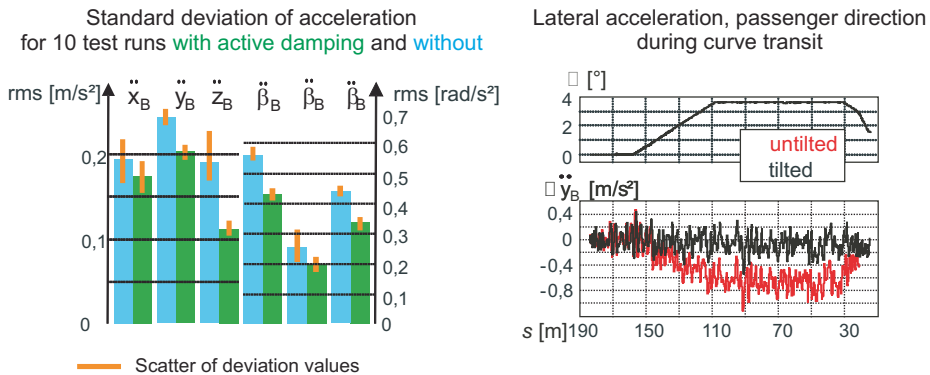


Figure 38: Measurement results for the active suspension.

The active suspension reduces the acceleration values in all degrees of freedom, in parts significantly. Compared to the passive suspension, the standard deviation of the acceleration could be reduced by 16% in lateral direction, 22% - 24% for rotation around the vehicle axes and even 41% for the lifting motion.

## Bibliography

- W. Beitz and G. Pahl. *Konstruktionslehre - Methoden und Anwendung*. Springer, Berlin, 4 edition, 1997.
- E. H. Bristol. On a new measure of interactions for multivariable process control. *IEEE Transactions on Automatic Control*, AC-11:133–134, 1966.
- R.M. Goodall and W. Kortüm. Active suspensions for railway vehicles - an avoidable luxury or an inevitable consequence? In *World congress on Railway Research*, pages 219–225, Cologne, Germany, 1990.
- H. Grotstollen. Design of long-stator linear motor drive for railcab test track. *Journal of Power Electronics*, 5:190–204, 2005.
- H. Grotstollen. Drive control and position measurement of railcab vehicles driven by linear motors. In *6th International Symposium on Instrumentation and Control Technology*, Beijing, China, 2006.
- M. Henke. *Antrieb mit doppeltgespeistem Linearmotor für ein spurgeführtes Bahnfahrzeug*. Number 533 in VDI-Fortschrittsberichte. VDI Verlag, Düsseldorf, Germany, 2003.
- M. Henke and H. Grotstollen. Modelling and control of a longstator-linearmotor for a mechatronic railway carriage. In *IFAC Conference on Mechatronic Systems*, Darmstadt, Germany, 2000.

- T. Hestermeyer. *Strukturierte Entwicklung der Informationsverarbeitung für die aktive Federung eines Schienenfahrzeugs*. Fahrzeugtechnik. Verlag Dr. Hut, Munich, Germany, 2006.
- T. Hestermeyer, O. Oberschelp, and H. Giese. Structured information processing for self-optimizing mechatronic systems. In *1st International Conference on Informatics in Control, Automation and Robotics (ICINCO)*, Setubal, Portugal, August 2004.
- U. Honekamp, R. Naumann, C. Homburg, and C. Sczyrba. Ein Konzept zur parallelen Echtzeitsimulation mechatronischer Systeme. In R. Fliieger and R. Grebe, editors, *Parallele Datenverarbeitung aktuell, Proceedings of the 6th Transputer-user-meeting*, Aachen, 1994.
- U. Honekamp, R. Stolpe, R. Naumann, and J. Lückel. Structuring approach for complex mechatronic systems. In *30th ISATA Conference on Mechatronics*, Florence, Italien, 1997.
- X. Liu-Henke. *Mechatronische Entwicklung der aktiven Feder-/Neigtechnik für das Schienenfahrzeug RailCab*. Number 589 in VDI Fortschritt-Berichte, Reihe 12: Verkehrstechnik/Fahrzeugtechnik. VDI Verlag, Düsseldorf, Germany, 2005.
- X. Liu-Henke and J. Ise. Objektorientierte Modellbildung, Simulation und Identifikation eines Feder-/Neigemoduls für ein Schienenfahrzeug. In *16. Symposium Simulationstechnik (ASIM'02)*, Rostock, 2002.
- X. Liu-Henke, J. Lückel, and K.-P. Jäker. Development of an active suspension/tilt system for a mechatronic railway carriage. In *1st IFAC Conference on Mechatronic Systems*, Darmstadt, 2000.
- X. Liu-Henke, J. Lückel, and K.-P. Jäker. An active suspension/tilt system for a mechatronic railway carriage. *IFAC - Control Engineering Practice*, 10:991–998, 2002.
- J. Lückel. The concept of mechatronic function modules applied to compound active suspension systems. Vortrag beim Symposium: *Research Issues in Automotive Integrated Chassis Control Systems*, International Symposium for Vehicle System Dynamic, 1992.
- J. Lückel. Das Systemkonzept: Neue Bahntechnik Paderborn (NBP). In *4. Int. Heinz Nixdorf Symposium*, Paderborn, 2000.
- J. Lückel. Mechatronik - Eine Einführung zur Vorlesung: Grundlagen der Regelungstechnik. Vorlesungsumdruck, MLaP - Universität Paderborn, 2004.
- J. Lückel, T. Hestermeyer, and X. Liu-Henke. Generalization of the cascade principle in view of a structured form of mechatronic systems. In *IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM 2001)*, volume 1, Como, Italy, 2001.
- J. Lückel, H. Grotstollen K. Jäker, M. Henke, and X. Liu. Mechatronic design of a modular railway carriage. In *IEEE/ASME Int. Conf. on Advanced Intelligent Mechatronics*, Atlanta, GA, USA, 1999.
- R. Naumann. *Modellierung und Verarbeitung vernetzter intelligenter mechatronischer Systeme*. Number 318 in VDI-Fortschrittsberichte, Reihe 20. VDI Verlag, Düsseldorf, 2000.
- O. Oberschelp, T. Hestermeyer, B. Kleinjohann, and L. Kleinjohann. Design of self-optimizing agent-based controllers. In *CfP Workshop 2002 - Agent-Based Simulation 3*, Passau, 7-9 April 2002.

- A. Pottharst. *Energieversorgung und Leittechnik einer Anlage mit Linearmotor getriebenen Bahnfahrzeugen*. Shaker Verlag, Aachen, 2006.
- A. Pottharst, K. Baptist, O. Schütze, J. Böcker, N. Fröhleke, and M. Dellnitz. Operating point assignment of a linear motor driven vehicle using multiobjective optimization methods. In *11th Intern. Power Electronics and Motion Control Conferences*, Riga, Latvia, 2004.
- A. Pottharst and H. Grotstollen. Radio based control of a linear motor for the nbp railway syste. In *5th International Symposium on Instrumentation and Control Technology*, Beijing, China, 2003.
- A. Pottharst and H. Grotstollen. Distributed control structure of the nbp test track with linear motor driven vehicles. In *16. Symposium Simulationstechnik (ASIM'02)*, Manaus, Brasil, 2005.
- T. Schneider, B. Schulz, C. Henke, and J. Böcker. Redundante Positionserfassung für ein spurgeführtes linearmotorgetriebenes Bahnfahrzeug. In *4. Paderborner Workshop Entwurf mechatronischer Systeme*, Paderborn, Germany, 2006.
- S. Skogestad and I. Postlethwaite. *Multivariable Feedback Control - Analysis and Design*. John Wiley & Sons, New York, 1996.
- R. Stolpe. *Verteilte kommunizierende mechatronische Funktionsmodule - Von der mechatronisch funktionalen Modularisierung bis zur verteilten HIL-Realisierung*. Number 393 in Reihe 20: Rechnerunterstützte Verfahren. VDI Verlag, Düsseldorf, 2004.
- R. Streiter, M. Boller, B. Riege, R. Schneider, and G. Himmelstein. Active lateral suspension for high speed trains - a step towards the mechatronic bogie. In *World Congress on Railway Research*, Cologne, 2001.
- S. Töpfer. *Die mechatronische Entwicklung des Parallelroboters TriPlanar*. Number 966 in Reihe 8: Me-, Steuerungs- und Regelungstechnik. VDI Verlag, Düsseldorf, 2002.
- B. Yang. *Regelung der Längs- und Nickbewegung des Linearmotors für ein vollaktives, spurgeführtes Bahnfahrzeug*. Shaker Verlag, Aachen, 2004.
- B. Yang and H. Grotstollen. Design of the power supply for the doubly-fed linear motor of railcab test track. In *Southern African Universities' Power Engineering Conference*, Cape Town, South Africa, 2007.

# Advanced Road Vehicles: Control Technologies, Driver Assistance

Dieter Ammon <sup>†</sup> and Werner Schiehlen <sup>‡</sup>

<sup>†</sup> DaimlerChrysler, Stuttgart, Germany

<sup>‡</sup> University of Stuttgart, Institute of Engineering and Computational Mechanics, Germany

**Abstract** Advanced road vehicles require sophisticated tools for the ride and handling analysis during the design phase. The corresponding model based development process is described, and driving performance measures are introduced. Active suspensions, active steering and the overall vehicle dynamics management are considered. Driver assistance systems are presented and discussed.

## 1 Introduction

Future developments of road vehicles include an interdisciplinary stimulation of the design concepts resulting in mechatronic driven cars. The Driver Assistance Systems are devices to control distance and to keep in lane on existing roads, relieving the driver and leading to a significant improvement in driving safety and comfort, Breuer (2007), Wallentowitz and Reif (2006).

The state-of-the-art will be shortly reviewed in this contribution. Starting with the entire vehicle ride and handling analysis, and driving performance measures will be introduced. The active components include active suspension, active steering and the related vehicle dynamics management. Finally some aspects of driver assistance systems are discussed.

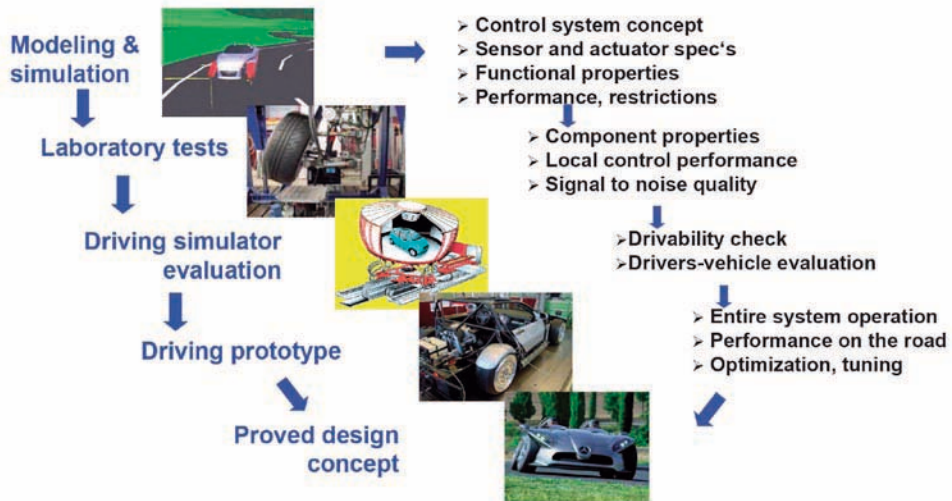
## 2 Entire Vehicle Ride and Handling Analysis

The development in vehicle dynamics are characterized by model-based methods, Ammon (1997) and Milliken (1995). The model type used is related to the design focus. Ride comfort simulations and durability load simulations are essential tools in the design process. Examples for modelling the components include a steering system and rubber parametrization. Further, the process proof model parametrization is considered.

Advanced vehicle require safety systems which are reliable up to the boundaries characterized by strong nonlinearities. On the other hand, the comfort systems demand high sensitivity with fine resolution of all the signals. The control systems require multi-disciplinary modelling including mechanics, hydraulics, electrics and electronics, Schiehlen and Schirle (2006), Rauh (2003). The entire vehicle development relies increasingly on simulation analysis. The engineers and managers are taking decisions more and more based on simulation results.



The model-based ride and handling development is shown in Figure 1. Starting with modeling one finally gets a proved design concept which includes also the computational and experimental testing of the components used.



**Figure 1.** Model-based ride and handling development

The model type applied depends on the design focus. The complexity of the model is characterized by its number of degrees of freedom and the frequencies of the motions involved. Figure 2 shows the model types of vehicle dynamics which are certainly below the structural vibrations with respect to complexity and frequency.

As an example, ride comfort simulation accuracy can be discussed by a vertical step input excitation of 10 mm height and 20 mm length with decreasing amplitude resulting in vertical and horizontal forces, Figures 3 and 4. Three different models are compared with one measurement. The longitudinal force shows higher frequencies than the vertical force but similar magnitude. Thus, for step inputs the longitudinal impact has a strong influence on the ride comfort, Ammon (2005b), Lugner and Plöchl (2005), Lugner et al. (2005).

More generally, ride comfort simulation has to consider

- the disturbance path: since road unevenness causes nonlinear 3-dimensional vibrations of the system tire – wheel – wheel-carrier – chassis,
- the vehicle driven excitations: since driving and drive train activities excite the entire chassis system, including tire & wheel, steering system, suspension,
- and the tire-driven excitations: i.e. unbalanced masses and self-excited shimmy.

Ride comfort analysis requires high accuracy in dynamic transfer path characteristics up to 25 Hz and high resolution road surface interfaces (road height, road-tire friction).

For the durability load simulation on obstacle with a constant height is used. The resulting tire forces are computed for a vertically fixed vehicle, Figure 5. Obviously, the vertical wheel load is the largest force.

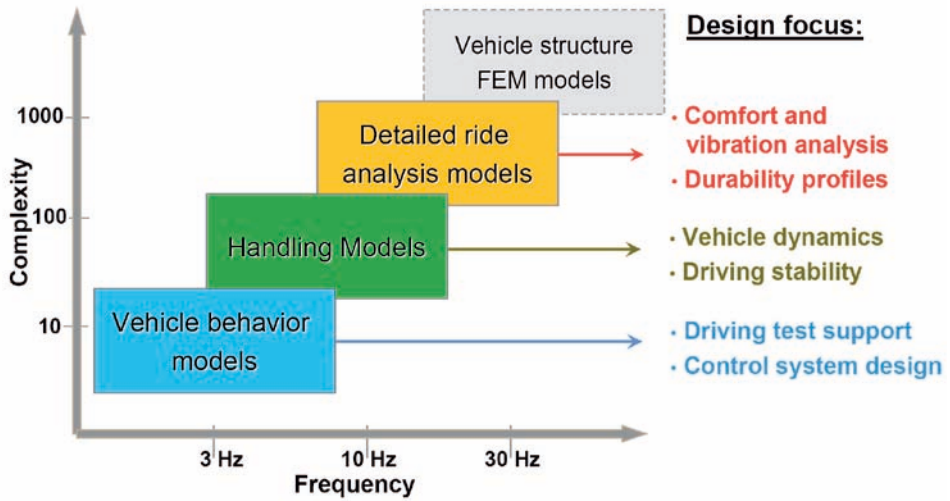


Figure 2. Design focus and related model type

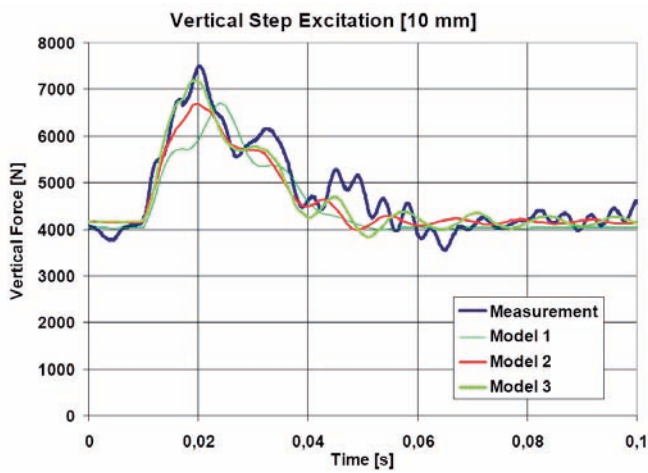


Figure 3. Ride comfort simulation, vertical

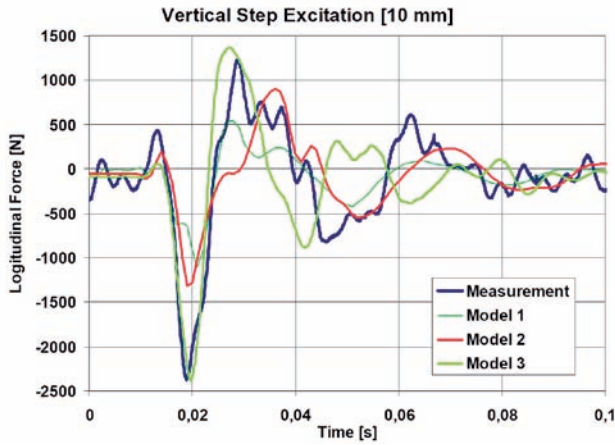


Figure 4. Ride comfort simulation, horizontal

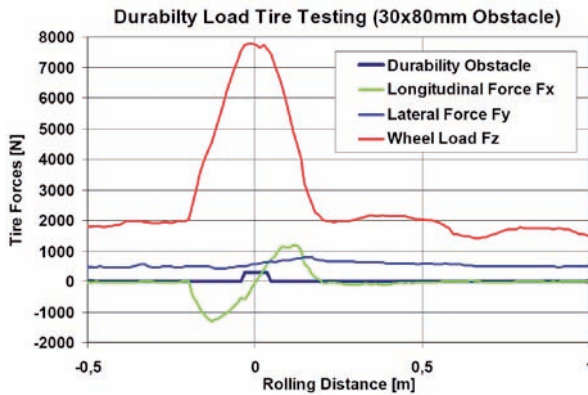


Figure 5. Durability load simulation

For handling the steering system is most important which is modelled as a four pole system with two inputs and outputs, respectively, subject to the vehicle speed and the pump speed, Groll et al. (2006). The block diagram of the steering is shown in Figure 6. A more detail representation of the steering system is shown in Figure 7. At the steering wheel angle  $\delta_{LR}$  and the steering torque  $M_{LR}$  are the inputs while the rack force  $F_{2S}$  and the rack displacements  $\Delta y_{2S}$  are acting on the steering linkage and the wheel itself.

On the other hand, for ride comfort the wheel suspension play a central role. A typical passenger car suspension is depicted in Figure 8 where damper and the rubber bushings are the critical components with respect to the modelling. The damper characteristics to be considered depend on frequency and magnitude of the vibrations as shown in Figure 9.

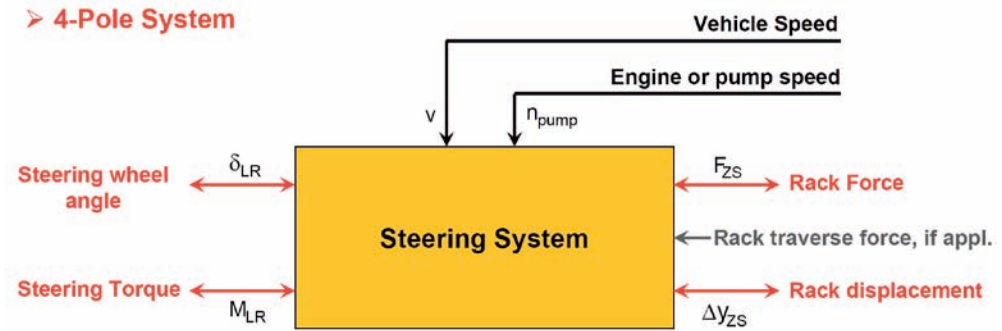


Figure 6. Component modelling: steering

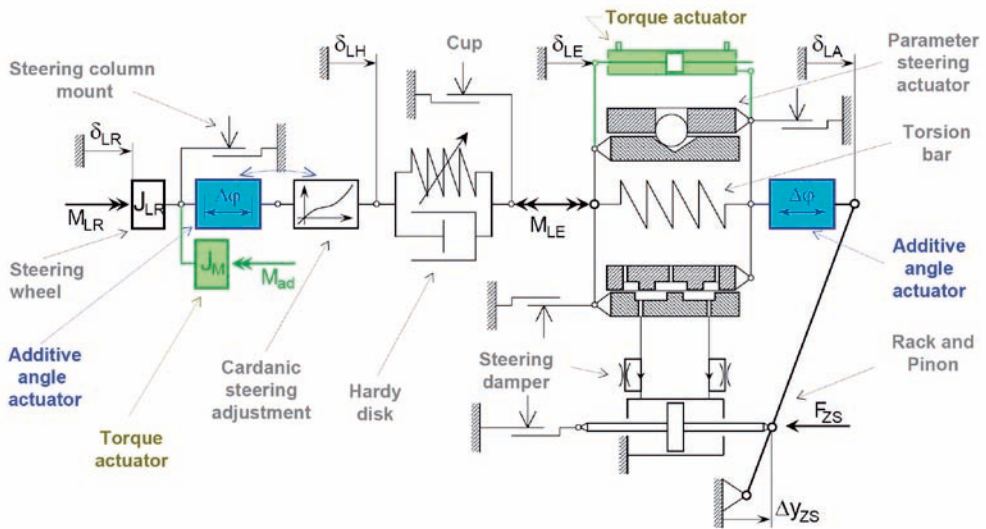


Figure 7. Active steering concept

First of all, dampers may be nonlinear and there are hysteresis loops due to friction and gas within the damper. The fluid flow within the damper may result in cavitation (in case of misuse) and the overall dynamic damper stiffness is frequency dependent, Matschinsky (1997).

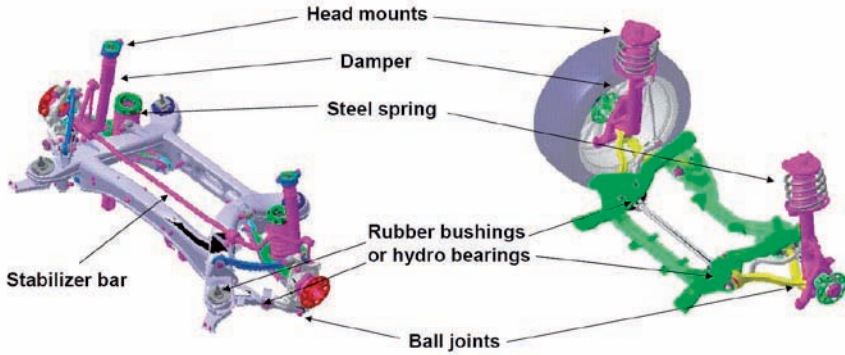


Figure 8. Wheel suspension of passenger car

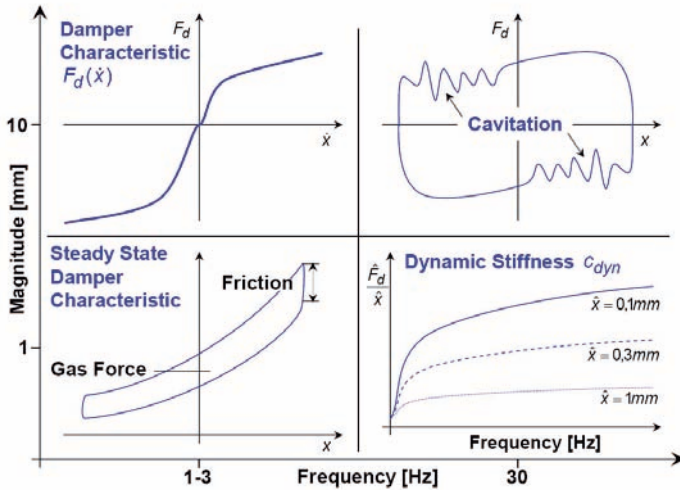


Figure 9. Damper characteristics depending on frequency and magnitude

The rubber bushings are characterized by their quasistatic and dynamic behaviour, Figure 10. Once again hysteresis loop can be observed indicating energy losses in the rubber moments. The dynamical stiffness and the related phase angle are frequency dependent, too. For the mathematical modeling fractional derivatives turned out to

be most appropriate, Figure 11, Sjöberg (2002). The frequency domain signal properties are represented by discrete time step filter models and the corresponding transfer characteristics.

The last step in the model process is the parametrisation task. A set of component model parameters has to be determined such that the simulated component properties approximate the real component behaviour within a well-known accuracy of error bandwidth.

The analysis tasks required are specified by the mechanical quantities motion dynamics, wheel loads, lateral slip, longitudinal slip and camber angle, Figure 12. The resulting forces and geometries have to be provided with a sufficient accuracy.

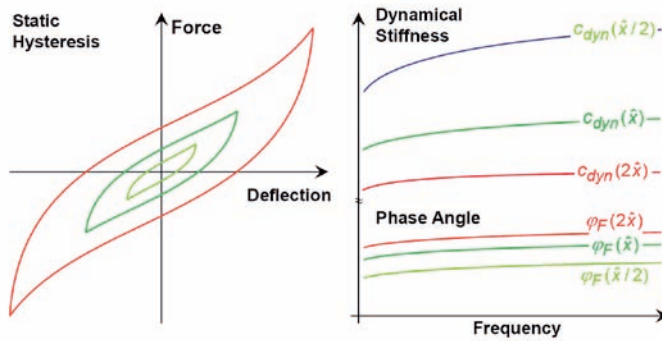


Figure 10. Rubber bushing characteristics

**Fractional derivatives modeling approach:**

**Definition:  $\alpha$ -times derivative of  $x$ :** 
$$y(t) = D^\alpha \{x(t)\} \equiv x^{(\alpha)}(t) = \frac{d^\alpha}{dt^\alpha} x(t) \quad \alpha \in \Re$$

**Frequency domain signal properties:** 
$$Y(\omega) = (j\omega)^\alpha X(\omega) = \omega^\alpha \left( \cos \frac{\alpha\pi}{2} + j \sin \frac{\alpha\pi}{2} \right) X(\omega)$$

**Discrete time steps filter model (moving average model):** 
$$y_k = D^\alpha \{x(t)\} = \lim_{n \rightarrow \infty} \sum_{l=0}^{n-1} a_l x_{k-l}$$

$$a_l = \frac{\Delta t^{-\alpha} \Gamma(l-\alpha)}{\Gamma(-\alpha) \Gamma(l+1)}$$

**Frequency domain transfer characteristics:**

$$F_{x \rightarrow y}^n(\omega) = \frac{Y}{X} = \sum_{k=0}^{n-1} (\cos k\omega\Delta t - j \sin k\omega\Delta t) a_k$$

Figure 11. Fractional derivatives modelling

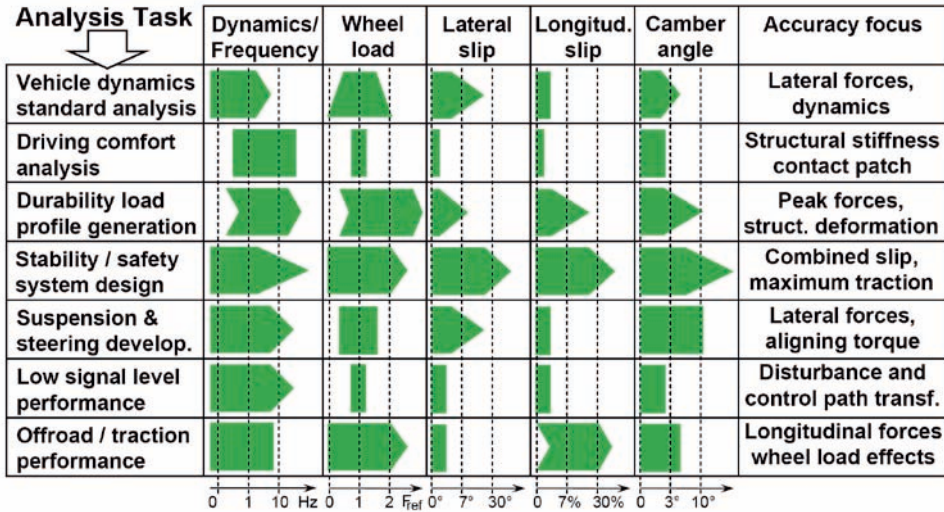


Figure 12. Requirements for model parametrization

### 3 Driving Performance Measures

Vehicle models are based on component models which result in geometrical and mechanical quantities. However, the driver is not a purely mechanical system and his sensation is directly related to the motions and forces acting on him. Therefore, the assessment of the vehicles by model-based methods requires human-specific driving performance measures. After discussion of the general goals the performance measures are introduced. Ride comfort evaluation and driving safety consideration is supplemented by a rollover analysis and some questions on “driving fun”.

The vehicle dynamics design and adjustment of present-day road vehicles are intended to ensure that

- the vehicle can be controlled with precision and is robust with regard to disturbances,
- the driver can master practically all driving situations,
- vehicle responses are predictable for the driver,
- the comfort, traction, stability and, if applicable, off-road characteristics appropriate to the vehicle type can be realized.

The performance measures related to the driver are driving comfort and driving safety. The driving comfort has to isolate the driver against vehicle or environment induced vibrations, respectively, and to enable fast and direct vehicle direction control to allow fun driving, too.

Driver and vehicle are operating in a closed control loop with disturbances from the environment as shown in Figure 13. The driving safety has to guarantee vehicle dynamics stability with appropriate damping, to guarantee stability of the entire control loop and to

enable fast and direct vehicle direction control. The driving safety is related to the wheel loads controlling the friction, and therefore also the suspension design does influence the driving safety.

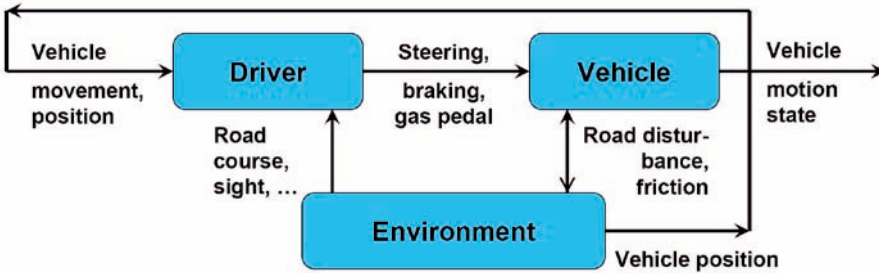


Figure 13. Driver-vehicle control loop

The ride comfort evaluation can be simply obtained on test roads by test drives. However, this procedure does not fit to the model-based design, Ammon et al. (2004). The seat contact acceleration data are evaluate by a vibration scheme resulting in more objective comfort measure. The human sensitivity is highly frequency dependent and requires a weighting of the accelerations as discussed by Schiehlen (2007), too. For the human perception assessment the standards VDI 2057 and ISO 2631 are available. As a matter of fact, the vibration sensitivity frequency responses has been recently updated in VDI 2057, Figure 14. The differences in the most important range for vehicle dynamics from 1 to 80 Hz are small so that the corresponding shape filter can be easily updated, too.

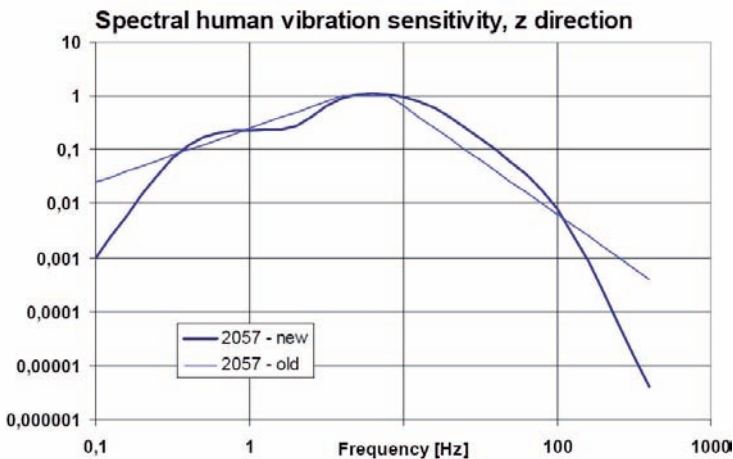
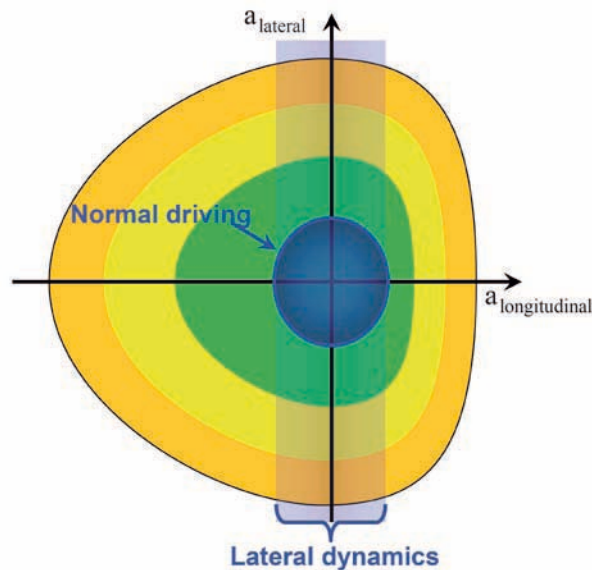


Figure 14. Standard vibration sensitivity by VDI 2057



The goals for driving safety include the basis vehicle dynamics design/setup, e.g., understeer gradient step-steer-response and limited yaw amplification. Open loop test of the vehicle without driver may be performed but most cases requires closed loop tests, Zomotor et al. (2003). The engineer has to fit vehicle reactions to driver expectations and to driver capabilities, too. However, the vehicles dynamics perception of the standard driver is related to normal driving situation at low acceleration levels. In this case, the vehicle reacts primarily linear. Only at higher acceleration levels, vehicle dynamics becomes extremely nonlinear, forces are limited, and delays are increasing.

Acceleration levels, in longitudinal and lateral direction are shown in Figure 15. It turns out that normal driving is mainly concerned with the limits in longitudinal direction.



**Figure 15.** Longitudinal and lateral accelerations

In addition to road test, driving simulator experiments are nowadays standard in vehicle dynamics testing, see Figure 16, Ammon (2005a) and Dragon (1996). The main advantage is that there is a low, nearly vanishing risk and a perfect repeatability of the experiments.

The rollover stability analysis is related to the global system driver-vehicle-environment as shown in Figure 17, Tseng et al. (2005) and Karnopp (2004). In the case of instability the actual vehicle reactions differ significantly to the drivers' expectations. As a result dangerous driving situations occur. For a basic rollover stability assessment the static stability factor (SSF) can be used, Figure 17, while the dynamic analysis is shown in Figure 18.

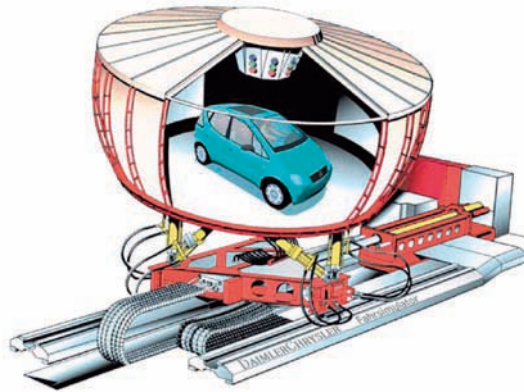


Figure 16. Driving simulator

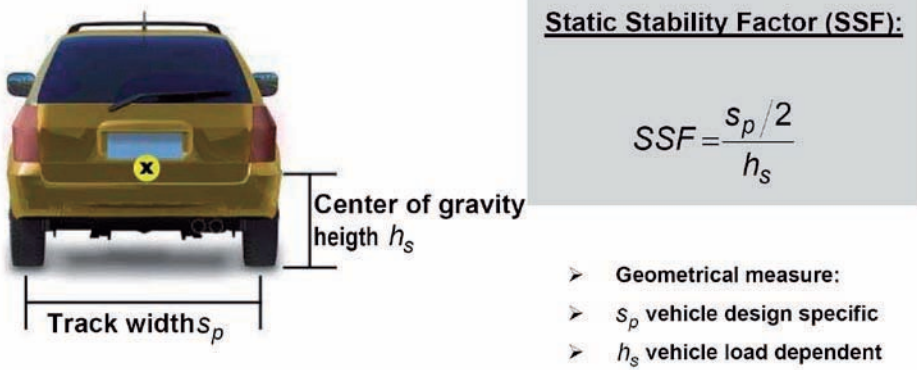
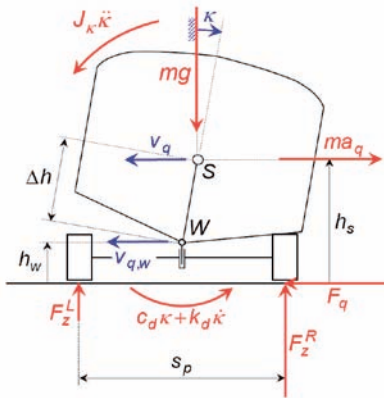


Figure 17. Static rollover stability analysis (source: NHTSA)

**Dynamic behavior:**



➤ Vehicle Body:  $ma_q = F_q$   $mg = F_z^L + F_z^R$

$$J_\kappa \ddot{\kappa} + \frac{s_p}{2} (F_z^R - F_z^L) = \Delta h mg \sin \kappa + h_s ma_q$$

➤ Roll center W moments:

$$J_\kappa \ddot{\kappa} + k_d \dot{\kappa} + c_d \kappa = \Delta h mg \sin \kappa + \Delta h ma_q \cos \kappa$$

$$k_d \dot{\kappa} + c_d \kappa + h_w F_q = \frac{s_p}{2} (F_z^R - F_z^L)$$

➤ for  $\kappa < 1$ :  $c_d \kappa - \Delta h mg \sin \kappa \approx c_\kappa \kappa$ ,  $\cos \kappa \approx 1$

$$J_\kappa \ddot{\kappa} + k_d \dot{\kappa} + c_\kappa \kappa = \Delta h ma_q \quad \text{with: } |\Delta F_z| \leq mg$$

or:

$$\Delta F_z = F_z^R - F_z^L = \frac{2}{s_p} (k_d \dot{\kappa} + c_d \kappa + h_w ma_q)$$

$$J_\kappa \ddot{\kappa} + \frac{s_p}{2} \Delta F_z - (c_d - c_\kappa) \kappa = h_s ma_q$$

**Figure 18.** Dynamic rollover stability analysis

From the wheel load spreading

$$\Delta f_Z = \frac{\Delta F_Z}{mg} \tag{3.1}$$

the efficient wheel load is found

$$\Delta f_Z^{eff} = \min \{ \max \{ \Delta f_Z; -1 \}; 1 \} . \tag{3.2}$$

Then, the dynamic rollover reserve value  $\rho$  reads as

$$\rho = 1 - \left| \Delta f_Z^{eff} \right| . \tag{3.3}$$

The driving fun depends on the handling behaviour of the vehicles as well as the driving situations chosen. Open questions are the customer expectations on the driving fun, and the market demands to fulfill the expectations.

## 4 Active Components and Vehicle Dynamics Management

Recent developments in vehicle dynamics are depicted in Figure 19 with respect to driving fun, driving safety and driving comfort. The latest developments like the research car F400, the rollover avoidance by ESP and active bogies are related to active components. Here active suspensions, active tire tilt control and active steering are considered, Ammon (2002), Rau and Ammon (2006).

An active suspension tries to overcome principal design conflicts between comfort and agility, see Figure 20. The design compatibility has to be moved as far as possible to the safety limit independently of the vehicle speed  $v$ . The active suspension principle

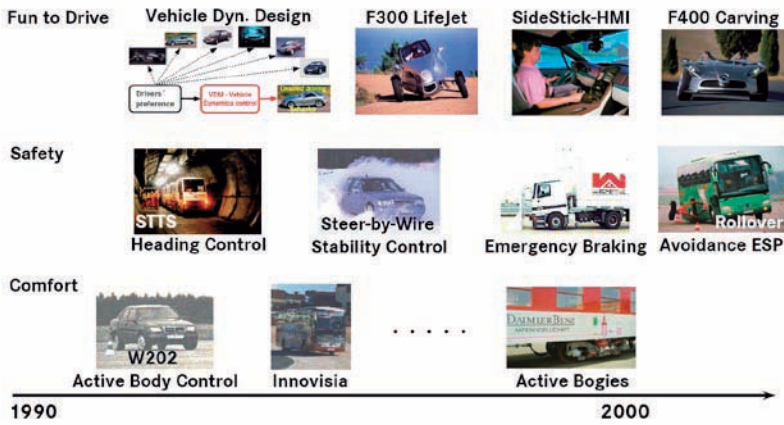


Figure 19. Developments in vehicle dynamics

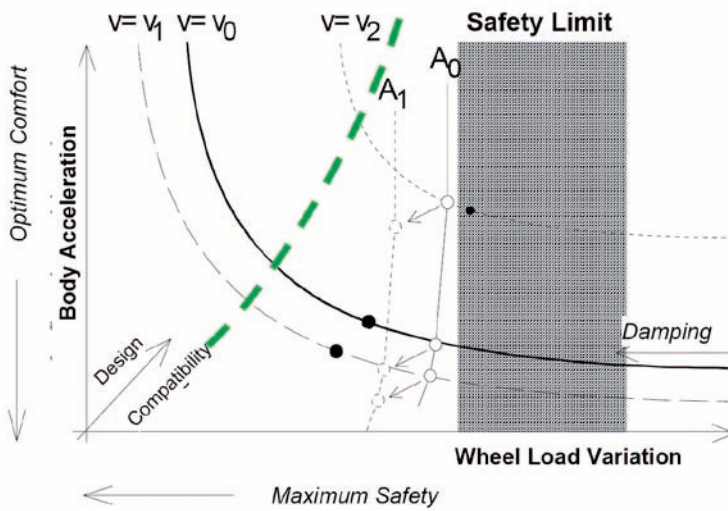


Figure 20. Active suspension design conflicts

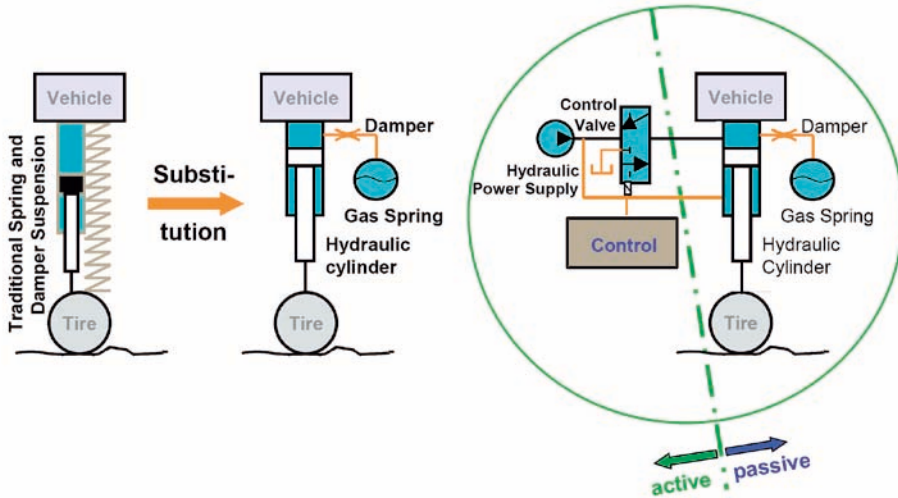


Figure 21. Active suspension components

is shown in Figure 21 and compared to passive suspensions. Realized versions of active suspension are the active hydro-pneumatic (AHP) by Citroen and the active body control (ABC) by DaimlerChrysler.

The active tire tilt control (ATTC) concept is shown in Figure 22. The basis concept combines the camber angle side forces known from motorbikes with specific rubber components as used in race cars. Thus, the side forces are increasing by 30 to 50 percent. The parts required for an ATTC are shown in Figure 23. In Ammon (2002) the design of the hydraulic components is discussed in detail.

The F400 carving ATTC system provides significantly increased traction resulting in enhanced vehicle dynamics. This properties can be used to

- build up a kind of lateral-ABS system
- enhance agility
- provide safety margins beyond usual limits
- emulate steer-by-wire functionalities

Active steering components are presented in Figure 24. There are two principle concepts

- additive angle
- additive torque

illustrated in Figure 25. The additive angle approach is used by BMW as superposition steering (Überlagerungslenkung) while the additive torque approach is realized by Volkswagen and by DaimlerChrysler in the A-Class. The additive torque enables a driver assistance system, and the additive angle enables subjective by vehicle agility improvements.

A vehicle dynamics management system is based on the possibility to design vehicle dynamics properties by software what is a unique feature of active systems found in

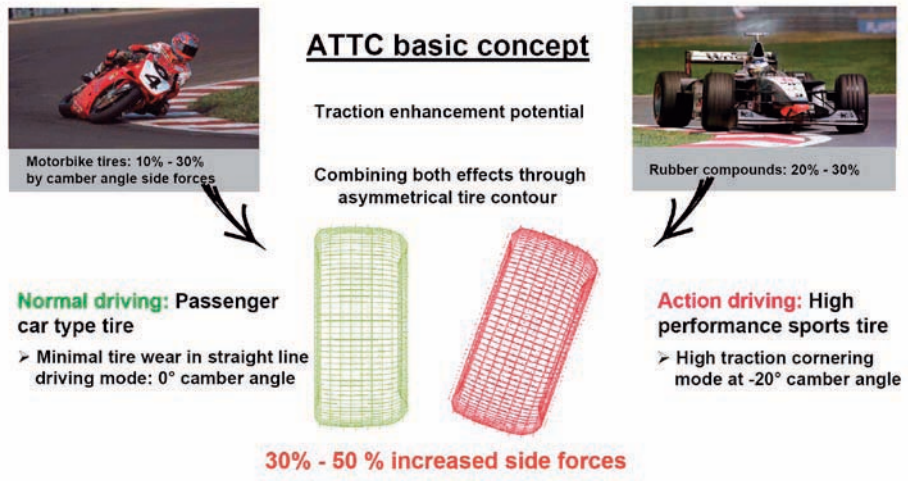


Figure 22. Active tire tilt control concept

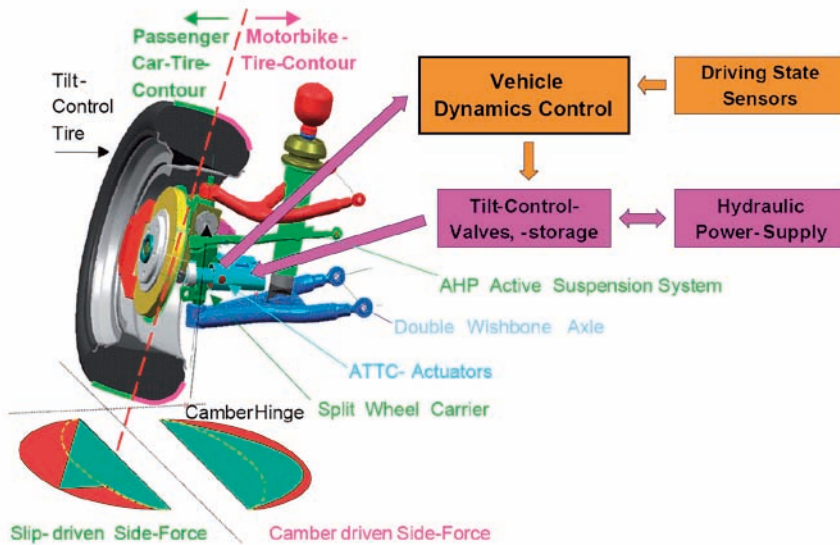


Figure 23. Active tire tilt control components

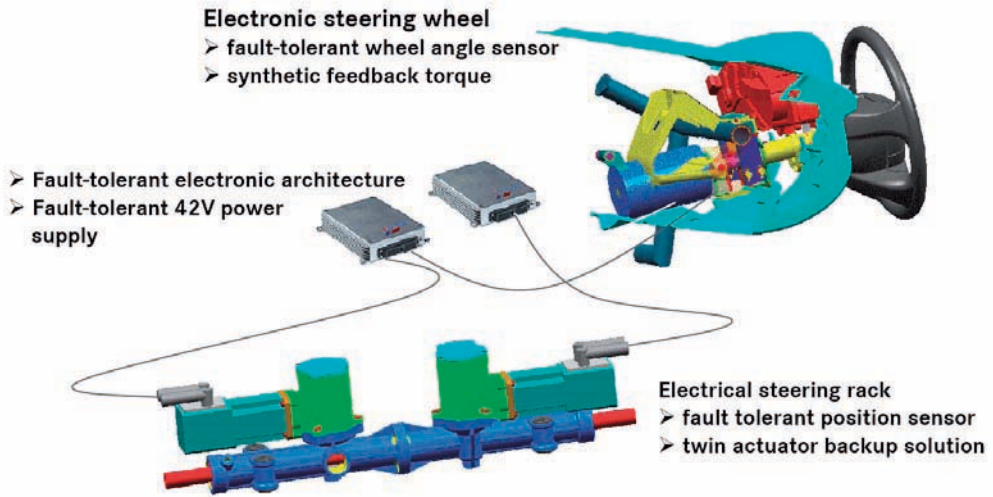


Figure 24. Active steering components

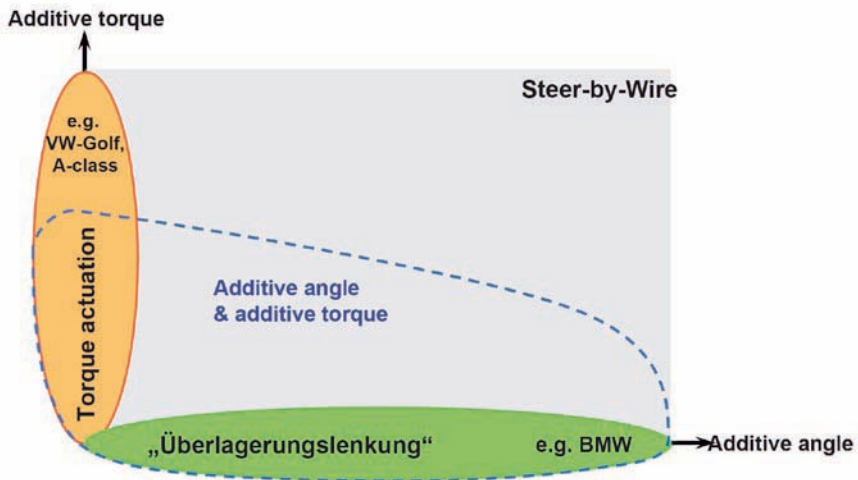


Figure 25. Active steering principle concepts

mechatronic vehicles. Figure 26 shows how the proven properties of a reference vehicle are transferred by a vehicle dynamics coordinator to a new designed mechanically different vehicle. Thus, the designer has less limitations with respect to shape, size space of a vehicle, and innovative vehicles with more benefits for the customer results in higher competitiveness on the global market.

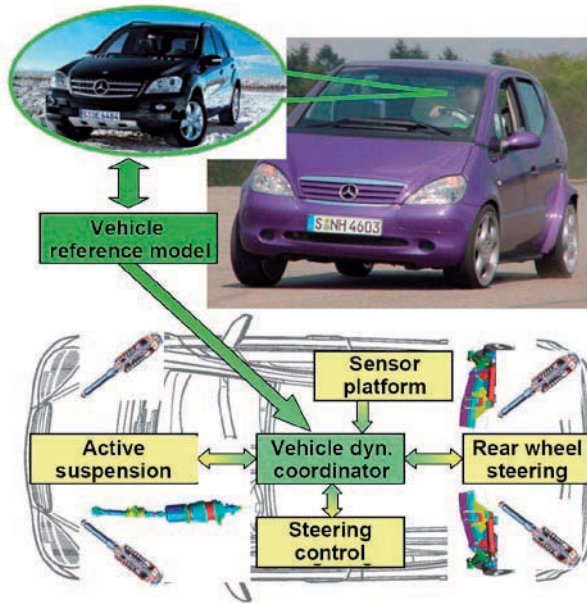


Figure 26. Vehicle dynamics management

## 5 Driver Assistance Systems

There are tasks can be made easier for the driver by assistance systems. The comfort assistance include

- Maneuvering Aid/Parking Support
- Distance Control Systems
- Low Speed Following
- Lane Keeping Support

The safety assistance is even more important, Breuer (2007). The following features are already available of planned for the near future:

- ESP Stability Control
- Brake Assistance
- Rollover Mitigation/Avoidance
- Optimal Driving Path Calculation
- Collision Mitigation/Collision Avoidance



- Integrated Safety Approaches

The ESP (Electronic Stability Program) takes responsibility for driving stability control. As shown in Figure 27 it took 36 year to bring ESP to the market. But then, within five years the number of driving accidents was reduced impressively by 30 to 40 percent. The Brake Assistant (BAS) solves the psychological problem of the hesitating driver, Figure 28. This driver assistance system was jointly developed by DaimlerChrysler and TWR/Lucas for passenger car.

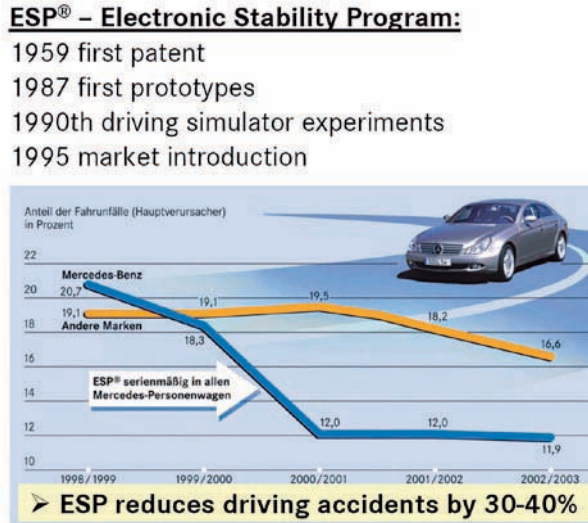


Figure 27. ESP driving stability control

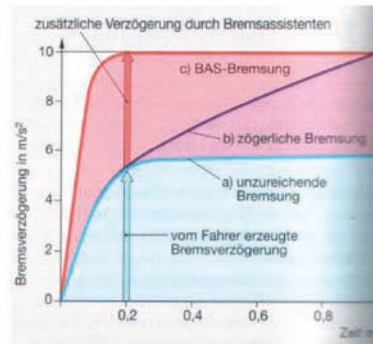
The Active Brake Assist (ABA) is used in Mercedes trucks. This feature is important since rear-end collisions are the most frequent type of heavy truck accidents. The system technology for the Active Brake Assist contribution to collision mitigation is shown in Figure 29. The components involved are a radar distance sensor with control unit interconnected by a data bus with the instrument panel and the conventional brake system.

The latest technology for integrated safety are precrash systems for passenger cars designed to reduce the damage caused by a collision. An example is the Mercedes Pre-Safe System, Figure 30.

This system provides brake support, tightens the seat belts, adjusts seat positions including rear seats and can also close any open windows and the sunroof if necessary. The operation of Mercedes Pre-Safe braking is depicted in Figure 31. The activation of the braking procedure may be initiated by the driver, by the system itself if required and by sensors immediately before the collision. In principle, the driver has the full responsibility, and the Pre-Safe option is completely restricted to emergency situations.

Problem:

- Numerous drivers hesitate in emergency braking situations
- They don't hit their brake pedal strong enough



without BAS (Stopping distance at 100 km/h)



with BAS: Reduction up to 45%



Figure 28. Brake Assist (BAS)

Active Brake Assist – System Technology

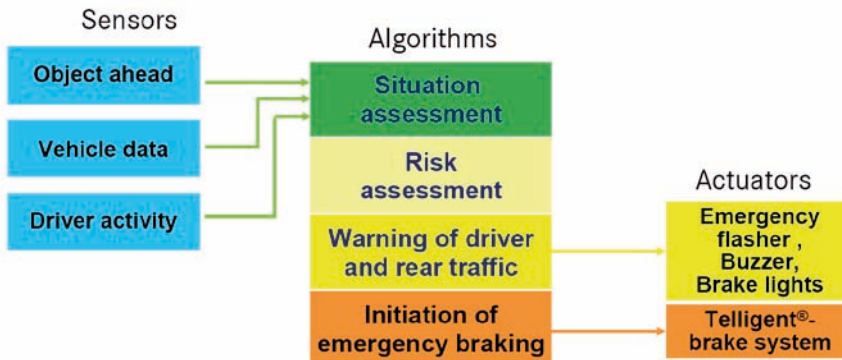


Figure 29. Collision mitigation by Active Brake Assist

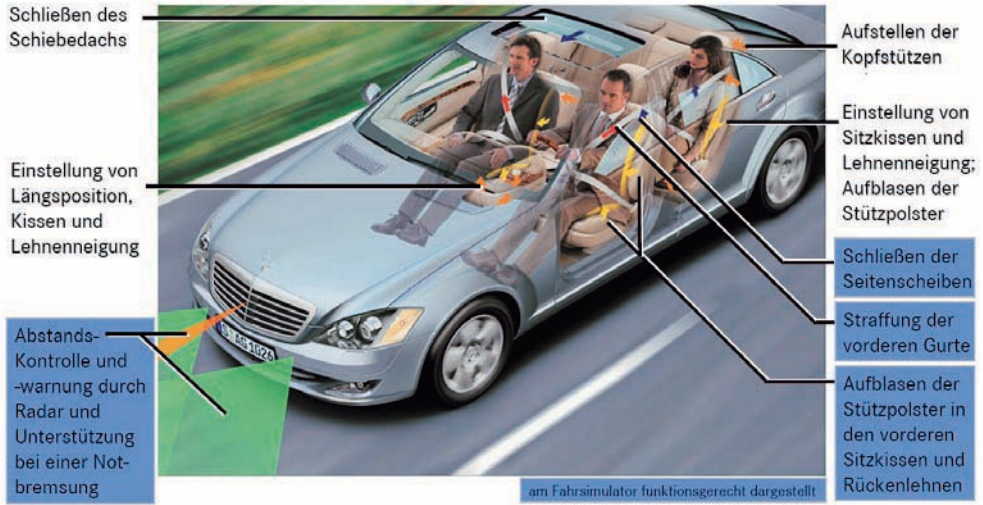


Figure 30. Integral safety: Mercedes Pre-Safe System

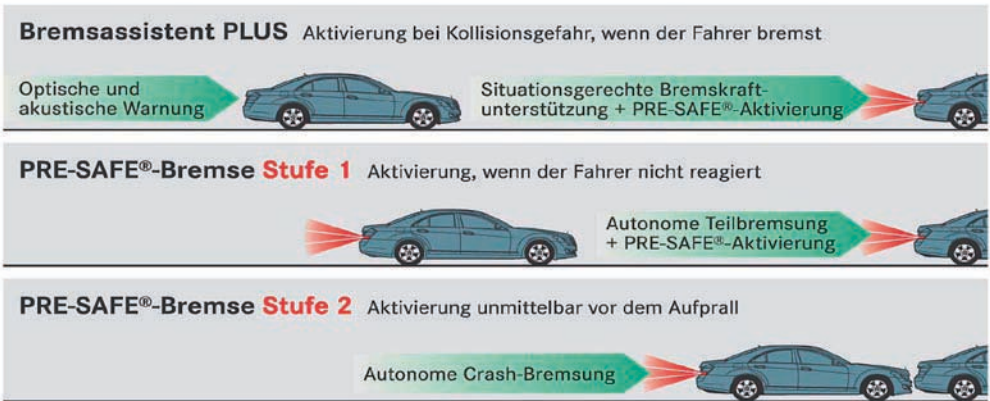


Figure 31. Mercedes Pre-Safe breaking

## 6 Conclusion

Today's models promise to approximate almost all functionally relevant vehicle dynamics properties. Expert users achieve high accuracy and simulation results reliability. However, for industrial standard applications only the field of classical vehicle dynamics analysis is well explored and process proved. It is still required to extend component and entire vehicle test rig capabilities towards higher frequencies/dynamics, higher loads resolutions, and to define parametrisation and validation standards - to achieve reliable results/quality. Active systems enable significant improvements of driving comfort, agility, and driving safety, especially with respect to situation-specific demands. Vehicle prototypes and simulation studies provide necessary informations to develop the most effective realization strategy.

## Bibliography

- D. Ammon. Fahrdynamiksimulation und Fahr simulatoren - Potenziale und Grenzen anspruchsvoller Entwicklungswerkzeuge. *Motion Simulator Conference*, pages 116–135, GZVB, Braunschweig, 2005a.
- D. Ammon. *Modellbildung und Systementwicklung in der Fahrzeugdynamik*. Teubner, Stuttgart, 1997.
- D. Ammon. New perspective in vehicle dynamics - the mechatronic suspension system of the F400 Carving. In *VDA-Kongreß Sicherheit durch Elektronik*, pages 345–354, 20.-21. März, 2002. VDA, Frankfurt, 2002.
- D. Ammon. Vehicle dynamics analysis tasks and related tyre simulation challenges. *Vehicle System Dynamics*, 43:30–47, 2005b.
- D. Ammon, H. Gimmler, K.-D. Hilf, J. Rauh, G. Scheible, and P. Stieß. Fahrzeugschwingungen - von der Fahrbahn anregung bis zum Komfortempfinden. *VDI-Tagung Humanschwingungen*, Darmstadt, 2004. VDI-Ber. Nr. 1821, pages 139–158, VDI-Verlag, Düsseldorf, 2004.
- J. Breuer. Driver Assistance Systems - From Cruise Control to Collision Mitigation. *VDA - Technischer Kongress 2007*, pages 199–207, Stuttgart. Henrich, Frankfurt, 2007.
- L. Dragon. FADYS handling simulations used in a real-time hardware-in-the-loop application with the ESP-controller. *AVEC 96 Int. Symposium on Advanced Vehicle Control*, 1:559–572, Aachen, 1996.
- M. Groll, S. Müller, T. Meister, and R. Tracht. Disturbance compensation with a torque controllable steering system. *Vehicle System Dynamics*, 44:327–338, No. 4, 2006.
- D. Karnopp. *Vehicle Stability*. Marcel Dekker, New York, 2004.
- P. Lugner, H. Pacejka, and M. Plöchl. Recent advances in tyre models and testing procedures. *Vehicle System Dynamics*, 43:413–436, No.6–7, Taylor & Francis, 2005.
- P. Lugner and M. Plöchl, editors. Tyre Models for Vehicle Dynamics Analysis. *Vehicle System Dynamics*, 43, Supplement, Taylor & Francis, 2005.
- W. Matschinsky. *Road Vehicle Suspensions*. Wiley Publishers, New York, 1997.
- W.F. Milliken and D.L. Milliken. *Race Car Vehicle Dynamics*. SAE, Warrendale, 1995.

- M. Rau and D. Ammon. Vernetzung von aktiven Fahrwerk-Regelsystemen am Beispiel Aktive Federung (ABC) und Fahrdynamik-Regelung (ESP). *5. Tag des Fahrwerks*, ika, Aachen, 2006.
- J. Rauh. Virtual development of ride and handling characteristics for advanced passenger cars. *Vehicle System Dynamics*, 40:135–155, 2003.
- W. Schiehlen. Vehicle and Guideway Modelling: Suspension Systems. *Published in this volume*, 2007.
- W. Schiehlen and T. Schirle. Modelling and simulation of hydraulic component of passenger cars. In S. Bruni and G. Mastinu, editors, *Proc. 19th Symp. Int. Ass. Veh. Syst. Dynamics*, pages 581–589. Taylor & Francis, London, 2006.
- M. Sjöberg. *On Dynamic Properties of Rubber Isolators*. PhD-Thesis, KTH, Stockholm, 2002.
- H.E. Tseng, J. Asgar, D. Hrovat, P. van der Jagt, A. Cherry, and S. Neads. Evasive manoeuvres with a steering robot. *Vehicle System Dynamics*, 43:199–216, No.3, 2005.
- H. Wallentowitz and K. Reif, editors. *Handbuch Kraftfahrzeugelektronik*. Vieweg, Wiesbaden, 2006.
- Z. Zomotor, D. Ammon, and D. Meljnikov. Modellbasierter Fahrversuch. In *VDI-Tagungsbericht Meß- und Versuchstechnik in der Fahrzeugentwicklung*, Stuttgart, 2003. VDI-Ber. Nr. 1755, pages 25–36, Düsseldorf, 2003.