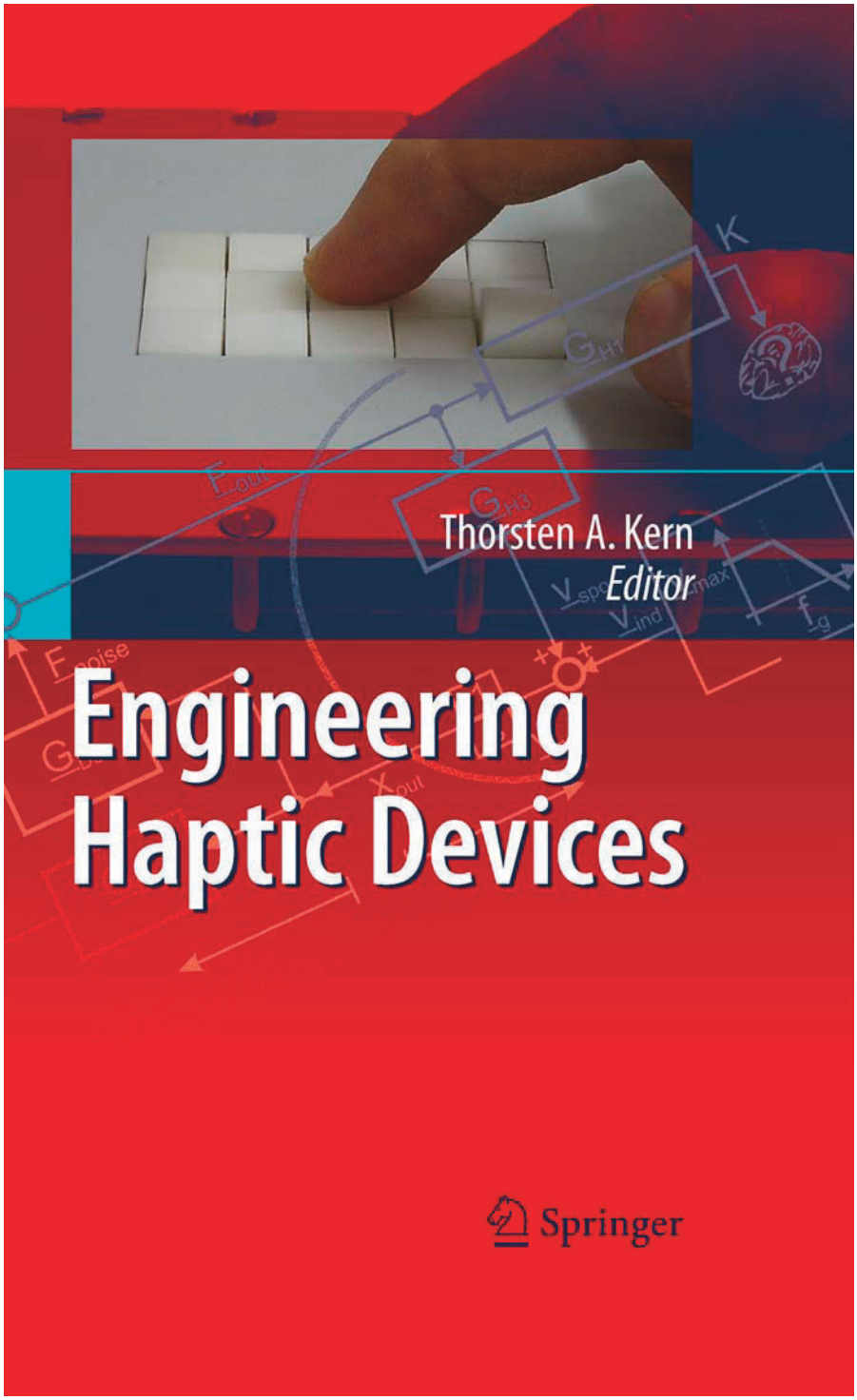

Thorsten A. Kern
Editor
Engineering
Haptic Devices
Springer

APPLE 1018

Engineering Haptic Devices

"*This page left intentionally blank.*"

Dr.-Ing. Thorsten A. Kern
Editor

Engineering Haptic Devices

A Beginner's Guide for Engineers

Springer

Editor
Dr.-Ing. Thorsten A. Kern
TU Darmstadt
Inst. Elektromechanische
Konstruktionen
Merckstr. 25
64283 Darmstadt
Germany

ISBN 978-3-540-88247-3 e-ISBN 978-3-540-88248-0
DOI 10.1007/978-3-540-88248-0
Springer Dordrecht Heidelberg London New York

Library of Congress Control Number: 2009930680

© Springer-Verlag Berlin Heidelberg 2009
This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilm or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.
The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Cover design: WMXDesign GmbH

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

The idea for this book was born in 2003. Originally intended as an addition to my dissertation, it was thought of filling a gap I had noticed: The regrettably small number of comprehensive recapitulating publications on haptics available for, e.g,. a technically interested person, confronted with the task of designing a haptic device for the first time. In 2004, besides a considerable number of conference proceedings, journals or PhD-theses, no document was available giving a summary of the major findings of this challenging subject. Thanks are due to the support provided by colleagues, especially by Prof. Dr.-Ing. Dr.-med Ronald Blechschmidt-Trapp and Dr.-Ing. Christoph Doerrer, enabling me to realize my plan in the following years. While writing my dissertation, I learnt that the areas to be covered by such a book would have to be much more extensive than originally expected. Nevertheless my mentor and supervisor Prof. Dr.-Ing. habil. Roland Werthschützky, to whom I have a special debt, encouraged me to pursue and finish this project during the time as a Post-Doc. It was funded by the DFG (KE1456/1-1) with special regard to the consolidation of our design methodology for haptic devices. Due to this funding the financial basis of this task was guaranteed.

The structuring of the themes made clear that the book would be considerably improved by contributions made by specialists of several areas. Thus began the multi-author project you are holding in your hands. Not only the authors explicitly named contributed to the book, but also former and current colleagues could be won over to a committed co-operation during the project, like Dr.-Ing. Markus Jungmann or Ingmar Stöhr, to name and thank at least two of them.

The first inquiry at Springer, our preferred publisher, resulted in an impressively positive reaction and consequently in a constructive co-operation with Dr. Christoph Baumann at any time. We owe to him the project's final configuration concerning language versions and date of publication. Let me add a word of gratitude to the many people involved in proofreading the book and removing major errors, all ahead of them Anika Kohlstedt, Sebastian Kassner, as well as to the numerous helpers of both family and friends.

For the English version of this book a professional language institute was originally hired which, however, was unable to provide the translation in time. Therefore it had to be done by the authors themselves. Most of the proofreading was done by Ursula and Jürgen Förnges both of whom I want to thank explicitly.

Without the help of all the people mentioned above and of many other unnamed supporters, this book would never have reached its present degree of completeness. I owe a great debt to all of them. A very special thanks is due to the *Institut für Elektromechanische Konstruktionen* at *Technische Universität Darmstadt*, Germany, for the extraordinary support given to me and many of my co-authors in our research and scientific work.

Thorsten A. Kern

"*This page left intentionally blank.*"

Preface

The term “haptics” unlike the terms “optics” or “acoustics” is not so well-known to the majority of people, not even to those who buy and use products related to haptics. The words “haptics” and “haptic” refer to everything concerning the sense of touch. “Haptics” is everything and everything is “haptic”, because it does not only describe pure mechanical interaction, but also includes thermal- and pain- (nociception) perception. The sense of touch makes it possible for humans and other living beings to perceive the “borders of their physical being”, i.e. to identify where their own body begins and where it ends. With regard to this aspect, the sense of touch is much more efficient than the sense of vision, as well in resolution as in the covered dihedral angle, e.g.: In the heat of a basketball match a light touch on our back immediately makes us aware of an attacking player we do not see. We notice the intensity of contact, the direction of the movement by a shear on our skin or a breeze moving our body hairs - all this is perceived without catching a glimpse of the opponent.

“Haptic systems” are divided into two classes[1]. There are the time-invariant systems (the keys of my keyboard), which generate a more or less unchanging haptic effect whether being pressed today or in a year’s time. Structures like surfaces, e.g. the wooden surface of my table, are also part of this group. These haptically interesting surfaces are often named “haptic textures”. Furthermore, there are active, reconfigurable systems, which change their haptic properties partly or totally dependent on a preselection - e.g. from a menu. Finally, there are combinations and hybrid forms of systems, which are presented and explained in the corresponding chapters. The focus of this book is on the technological design criteria for active, reconfigurable systems, providing a haptic coupling of user and object in a mainly mechanical un-

[1] In engineering there are three terms which are often used but do not have a definite meaning: System, Device and Component. Systems are - depending on the task of the designer - either a device or a component. A motor is a component of a car, but for the developer of the motor it is a device, which is assembled from components (spark-plug, cocks, knocking-sensor). It can be helpful when reading a technological text to replace each term with the word “thing”. Although this suggestion is not completely serious, it surprisingly increases the comprehensibility of technical texts.

derstanding. Thermal and nociceptive perceptions are mentioned according to their significance but are not seriously discussed. This is also the case with regard to passive haptic systems.

The fact that you have bought this book suggests that you are interested in haptics. You might have already tried to sketch a technical system meant to fool haptic perception. And this attempt may have been more or less successful, e.g. concerning your choice of the actuators. Maybe, you are just planning a project as part of your studies or as a commercial product aimed at improving a certain manual control or at introducing a new control concept. Approaches of this kind are quite frequent. Many of the first active haptic systems were used in airplanes, to make aware of critical situations by a vibrating control handle. Nowadays, the most wide-spread active haptic system surely is the vibration of a mobile-phone. It enables its user to notice the reception of a message without visual or auditory contact, whereby even the type of the message - SMS or phone call - is coded in this buzzing haptic signal. More complex haptic systems can be found in automotive technology, as e.g. reconfigurable haptic control knobs. They are typically located in the center of the control console and are usually part of complex luxury limousines. Today, multidimensional haptic interaction is no longer limited to navigation- or modeling purposes of professional users, but has also found its way into interaction during computer gaming. Maybe, you are a member of the popular group of doctors and surgeons actively using haptics in medical technology. There has been a continuous increase of, the complexity of the tools for minimally-invasive surgery - longitudinal instruments with a limited degree of freedom to inspect and manipulate human tissue through small artificial or natural openings in the human body. This automatically results in the loss of the direct contact between surgeon and the manipulated tissue. For decades, the wish to improve the haptic feedback during such kinds of applications and/or the realization of training methods for minimally-invasive surgery has been a high motivation for researchers in haptic device design, however without a satisfactory commercial breakthrough, yet significant improvements in telemanipulation and simulation have been achieved.

Despite of or even because of the great variety of projects in industry and research working with haptic systems, the common understanding of “haptics” and the terms directly referring to it, like “kinaesthetic” and “tactile” are by no means as unambiguous and indisputable as they should be. In this book, we, the authors, intend to offer you a help to act more safely in the area of designing haptic devices. This book will begin with the presentation of the terminology and its usage according to what we regard as appropriate. Then it will provide a deeper understanding of haptics and a simplified engineering description, and will finally lead to concrete instructions and recommendations for the design of technologically complex haptic systems.

Besides the intention to generate real hardware design, there is another reason for dealing with haptic device design: A continuing ambition to extend one’s knowledge of haptic perception. This discipline, named “psychophysics” is an “unsharp”, non deterministic science formulating hypotheses and systematically checking them with the help of experiments and observations. These experiments are paramount to any progress. Consequently, special attention has to be paid to their quality and the

parameters observed. As a by-product of this important science of haptic research a plurality of devices and technical systems have been built. In fact psychophysics uses expertise in many different disciplines to solve its problems. As a consequence, important and creative engineers and scientists like Prof. HONG TAN and Prof. VINCENT HAYWARD have not only been designing high fidelity and very efficient haptic devices, but are also heavily involved in the research on psychophysical parameters. Psychophysics with emphasis on haptic questions is a very dynamic science. Every year, an uncounted number of results and experiments are published at congresses and in journals. Lately, MARTIN GRUNWALD [79] has published a notable summary of the latest state of knowledge. The book you are holding in your hands does not claim to keep up with every detail of this psychophysical progress. However, it tries to include as many of its findings as possible into the design of haptic devices. This book has been written by and is addressed to engineers of all the disciplines mentioned before: Design-engineers representing mechanical engineering, hardware-near electrical engineering, control-engineering, software-engineering or as a synergy of expertise in all disciplines of mechatronics.

As said before, the haptic sense is doubtlessly gaining in importance. This can be concluded from the great number of scientific publications on this subject and from the fact that all relevant distal senses like the senses of sight and hearing have already been provided with synthetic information in almost perfect quality in every-day life. "Perfect quality" may have different meanings depending on the actual context. A realistic rendering of a sensual experience can be an important requirement. The resolution of a 3D-monitor has to be below the resolution capability of the human eye in color dynamics and spatial distances between the picture elements (pixels). Sounds have to be traceable in space and must not interfere with artifacts of the storage- or transmission medium. In different circumstances attracting attention can be another "perfect quality". Typically, warning signals in the dashboard of a car are visual examples, so are acoustic signals in the cockpit of an airplane. Another demand on "perfect quality" can be the simultaneous requirement of high discriminabilty and large range - just think of navigational signals for ships. Both areas - optics and acoustics - have been subject to intense research for decades and have been provided with numerous intelligent device designs. In many cases the borders of the human capability of perception of the information provided have been reached or even crossed nowadays. At this point it is obvious to make use of another human sense to transmit information. Another motivation is the true-to-life simulation of virtual environments. After visual and auditory presentation having reached a high quality, the focus is directed to the haptic sense as being the next important one. Only this sense enables us to experience our physical borders and the synergy of interaction and perception.

Further areas of haptic research are telepresence and telemanipulation systems. In these cases, an intuitive and immediate feedback is a prerequisite for a safe handling of e.g. dangerous and / or valuable materials. There are reasons enough and to spare for dealing with the design of haptic devices which are demanded by the market. However, experts are rare and the access to this subject is difficult. The design of haptic devices demands interdisciplinary knowledge which should include

the basics of the properties of haptic perception and its dynamic-dependence on amplitude and frequency. Furthermore an overview of technological solutions, like the designs of actuators, kinematics or complete systems including software-solutions and the interfaces to simulations and virtual reality systems may be extremely helpful. For designing virtual reality systems it is also necessary to know the concepts of haptic renderings to enhance communication between soft- and hardware engineers.

The authors of this book regard their task as being fulfilled as soon as this book helps to fascinate more design-engineers by the development of haptic devices, thus speeding up the creation of more and better haptic systems available on the market.

Darmstadt, February 2009 *Thorsten A. Kern*



Thorsten A. Kern received his Doctor in electrical engineering in 2006 at the Technische Universität Darmstadt for the design of a haptic assistive system for cardio-catheterizations. He is working as a group leader at Continental Corporation, and is responsible for the development and production of actuators for visual and haptic applications. Before he was working in the area of medical simulation technology and VR-applications using haptics at PolyDimensions GmbH. His research focused on engineering methods to quantify the haptic impression of actuators and devices. He is associated lecturer for biomedical engineering (*t.kern@hapticdevices.eu*).



Marc Matysek graduated in electrical engineering at the Technische Universität Darmstadt in 2003. There, he is working as a PhD student at the Institut für Elektromechanische Konstruktionen since 2003. His research focuses on the technology of dielectric elastomer actuators and their use in applications as tactile displays (*m.matysek@hapticdevices.eu*).



Oliver Meckel received his degree for mechanical engineering in 2002 at Technische Universität Darmstadt. He is working as group leader of the technical department at the Wittenstein motion control GmbH. He is responsible for the development, simulation, testing, and technical product support of electrical gear motors and actuator systems. Before he was working at the Institute for Flight Systems and Control at Technische Universität Darmstadt. His research was focused on adaptive control systems and the development of Unmanned Autonomous Vehicles (*o.meckel@hapticdevices.eu*).



Jacqueline Rausch received her diploma in electrical engineering and information technology in 2005 at the Technische Universität Darmstadt. There, she is currently working as a research associate at the Institute for Electromechanical Design. Her main topic is the development of miniaturized piezoresistive multi-axis force/torque sensors to be integrated at the tip of minimally invasive surgical manipulators (*j.rausch@hapticdevices.eu*).



Photo: Lichtbildatelier Darmstadt

Alexander Rettig received his diploma in mathematics at the Technical Universität Darmstadt in 1998. He is working as team leader software development for Poly-Dimensons GmbH in the area of medical graphics and simulation applications. Before he was research associate at the Fraunhofer Institute for Graphics Darmstadt, where he focused on the integration of haptics into virtual reality systems and development of virtual reality applications (*a.rettig@hapticdevices.eu*).



Andreas Röse received his engineering degree in 2002 in electromechanical engineering. He is working as a research associate at the Institute of Electromechanical Design at the Technische Universität Darmstadt. His research work is focussed on parallel-kinematic mechanisms for minimally invasive surgical telemanipulators including numerical calculation techniques, mechanical design and performance evaluation methods. From 2002 to 2005 he was with Fresenis Medical Care and developed hemodialysis machines with emphasis on safety systems and control architecture (*a.roese@hapticdevices.eu*).



Stephanie Sindlinger b. Klages received her diploma in electrical engineering and information technology at the Technische Universität Darmstadt in 2004. She is working as a research associate at the Institut für Elektromechanische Konstruktionen. There she develops a haptic device for cardiologic catheterisations. Her expertises are the development of piezoelectric ultrasonic actuators and the design of the user-interface for integrating the system in the operational environment (*s.sindlinger@hapticdevices.eu*).

"*This page left intentionally blank.*"

Contents

Symbols xxi

Part I Some Basics of Haptics

1 Motivation and Application of Haptic Systems 5
1.1 The Meaning of Haptics from a Philosophical and Social Viewpoint 5
1.1.1 Haptics as a *Physical Being's Boundary* 5
1.1.2 Formation of the Sense of Touch 8
1.1.3 Special Aspects of the Design Process 11
1.2 The Significance of Haptics in Everyday Professional Life 13
1.2.1 The Sense of Touch in Everyday Medical Life 13
1.2.2 The Sense of Touch in the Cockpit 14
1.2.3 The Sense of Touch at the Desk 16
1.2.4 The Sense of Touch in Music 17

2 Terminology 19
2.1 Scientific Disciplines as Part of Haptic Research 19
2.2 Terms and Terminology Used for the Description of Haptic Systems 21
2.2.1 Basic Concepts of Haptics 21
2.2.2 Definition of Haptic Systems 22
2.2.3 Parameters of Haptic Systems 26
2.2.4 Characterization of Haptic Object Properties 28
2.2.5 Technical examples 29

3 Biological Basics of Haptic Perception 35
3.1 The Sense of Touch and its Biology 35
3.2 Haptic Perception 39
3.2.1 Psychophysical Concepts 40
3.2.2 Frequency Dependency 46
3.2.3 Characteristics of Haptic Interaction 50
3.3 Conclusions from the Biology of Haptics 54

3.3.1 Stiffnesses 54
3.3.2 One Kilohertz - Significance for the Mechanical Design? . . . 56

4 Modeling the User 59
4.1 Mapping of Frequency Ranges onto the User's Mechanical Model . . 59
4.2 A Model of the User as a Mechanical Load 63
4.2.1 Types of Grasps 64
4.2.2 Measurement Setup and Equipment 66
4.2.3 Models 68
4.2.4 Modeling Parameters 69
4.3 Modelling Haptic Perception 80
4.4 Application Remarks 84
4.4.1 Kinaesthetic System 84
4.4.2 Tactile System 85
4.4.3 Examples in Time and Frequency Domain 87
4.5 Summarizing Remarks on the Application of the Method 93

5 Internal Structure of Haptic Systems 95
5.1 System Engineering 95
5.1.1 Open-Loop Impedance Controlled 96
5.1.2 Closed-Loop Impedance Controlled 97
5.1.3 Open-Loop Admittance Controlled 98
5.1.4 Closed-Loop Admittance Controlled Devices 100
5.1.5 Qualitative Comparison of the Internal Structures of Haptic Systems 102
5.2 Abstraction of the System Components 104
5.2.1 Basic System Analysis 104
5.2.2 Analysis of System Components 105

Part II Designing Haptic Systems

6 Identification of Requirements 115
6.1 The Right Questions to Ask 115
6.1.1 Interaction as a Classification Criterion 117
6.1.2 Cluster "Kinaesthetic" 118
6.1.3 Cluster "Surface-Tactile" 119
6.1.4 Cluster "Vibro-Tactile" 120
6.1.5 Cluster "Vibro-Directional" 121
6.1.6 Cluster "Omni-Dimensional" 121
6.1.7 Cluster "always" 122
6.2 Experiments Together With the Customer 122
6.3 Requirement Specifications of a Haptic System 124
6.4 Order of Technical Decisions 128

7 Control of Haptic Systems 131
7.1 Introduction 131
7.2 System Description 132
7.2.1 Linear SISO-System Description 132
7.2.2 Linear State Space Description 136
7.2.3 Nonlinear System Description 139
7.3 System Stability 141
7.3.1 Analysis of Linear System Stability 141
7.3.2 Analysis of Nonlinear System Stability 145
7.3.3 Passivity in Dynamic Systems 150
7.4 Control Design for Haptic Systems 152
7.4.1 Structuring of the Control Design 152
7.4.2 Requirement Definition 154
7.4.3 Control Law Design 156
7.5 Conclusion 163

8 Kinematic Design 165
8.1 Basics 165
8.1.1 Mechanisms and their Classification 166
8.1.2 Calculating Kinematics 168
8.1.3 Transfer Characteristics and Jacobian Matrix 170
8.1.4 Optimizing the Transfer Characteristics 174
8.2 Serial Mechanisms 175
8.2.1 Topological Synthesis 175
8.2.2 Calculation of the Kinematic Problems 177
8.2.3 Example of a Serial Mechanism 179
8.3 Parallel Mechanisms 182
8.3.1 Topological Synthesis 182
8.3.2 Calculation of the Kinematic Transfer Functions 185
8.3.3 Examples of a Parallel Mechanism 186
8.4 The Complete Process of Kinematic Design 188

9 Actuator Design 191
9.1 General Facts about Actuator Design 192
9.1.1 Overview of Actuator Principles 192
9.1.2 Actuator Selection Aid Based on its Dynamics 196
9.1.3 Gears 197
9.2 Electrodynamic Actuators 199
9.2.1 The Electrodynamic Effect and its Influencing Variables . . . 199
9.2.2 Actual Actuator Design 213
9.2.3 Actuator Electronics 218
9.2.4 Examples for Electrodynamic Actuators in Haptic Devices . 224
9.2.5 Conclusion about the Design of Electrodynamic Actuators . . 227
9.3 Electromagnetic Actuators 229
9.3.1 Magnetic Energy 229

9.3.2 Design of Magnetic Circuits 232
9.3.3 Examples for Electromagnetic Actuators 236
9.3.4 Magnetic Actuators in Haptic Devices 239
9.3.5 Conclusion about the Design of Magnetic Actuators 241
9.4 Piezoelectric Actuators 243
9.4.1 The Piezoelectric Effect 243
9.4.2 Designs and Properties of Piezoelectric Actuators 247
9.4.3 Design of Piezoelectric Actuators for Haptic Systems 258
9.4.4 Procedure for the Design of Piezoelectric Actuators 258
9.4.5 Piezoelectric Actuators in Haptic Systems 266
9.5 Electrostatic Actuators 277
9.5.1 Definition of the Electric Field 277
9.5.2 Designs of Capacitive Actuators with Air-Gap 279
9.5.3 Dielectric Elastomer Actuators 285
9.5.4 Designs of Dielectric Elastomer Actuators 289
9.5.5 Electro-Rheological Fluids 292
9.6 Special Designs of Haptic Actuators 301
9.6.1 Haptic-Kinaesthetic Devices 301
9.6.2 Haptic-Tactile Devices 305

10 Force Sensor Design 313
10.1 Constraints 313
10.1.1 Topology of the Device 313
10.1.2 Contact Situation 314
10.1.3 Mechanical Properties of Measuring Objects 315
10.1.4 Texture of Measuring Objects 317
10.1.5 Selection of Design Criteria 319
10.2 Sensing Principles 321
10.2.1 Basics of Elasto-mechanics 322
10.2.2 Resistive Strain Measurement 325
10.2.3 Piezoresistive Silicon Sensoren 329
10.2.4 Further Resistive Sensors 334
10.2.5 Capacitive Sensors 335
10.2.6 Optic Sensors 340
10.2.7 Piezoelectric Sensors 347
10.2.8 Exotic Ones 350
10.3 Selection of a Suitable Sensor 352

11 Application of Positioning Sensors 357
11.1 Basic Principles of Position Measurement 357
11.1.1 Incremental Principle 357
11.1.2 Absolute Measurement Principle 358
11.2 Requirements in the Context of Haptics 359
11.3 Optical Sensors 360
11.4 Magnetic Sensors 364

11.5 Other Displacement Sensors 365
11.6 Electronics for Absolute Positions Sensors 367
11.7 Acceleration and Velocity Measurement 368
11.7.1 Integration and Differentiation of Signals 369
11.7.2 Induction as a Velocity Measure 370
11.7.3 Force Sensors as Acceleration Sensors 371
11.8 Conclusion on Position Measurement 371

12 Interface Selection 373
12.1 Border Frequency of the Transmission Chain 374
12.1.1 Bandwidth in a Telemanipulation System 374
12.1.2 Bandwidth in a Simulator-System 374
12.1.3 Data Rates and Latencies 376
12.2 Concepts for Bandwidth Reduction 377
12.2.1 Analysis of the Required Dynamics 377
12.2.2 Local Haptic Model in the Controller 377
12.2.3 Event-Based Haptics 378
12.2.4 Movement Extrapolation 378
12.2.5 Compensation of Extreme Dead Times 379
12.2.6 Compression 380
12.3 Technical Standard Interfaces 381
12.4 Final Remarks about Interface Technology 385

13 Software Design 387
13.1 Overview About the Subject “Virtual Reality” 388
13.2 Design and Architecture of VR-Systems 391
13.2.1 Hardware Components 391
13.2.2 Device Integration and Device Abstraction 392
13.2.3 Software Components 394
13.2.4 Simulation 396
13.2.5 Subsystems for Rendering 399
13.2.6 Decoupling of the Haptic Renderer from other Sense Modalities 401
13.2.7 Haptic Interaction Metaphors 403
13.3 Algorithms 405
13.3.1 Virtual Wall 406
13.3.2 “Penalty” Methods 409
13.3.3 Constraint-based Methods 412
13.3.4 6 DoF-interaction: Voxmap-PointShell-Algorithm 415
13.4 Collision Detection 421
13.5 Conclusion 429

14 Final Remarks on the Design of Haptic Systems 431

Part III Appendix

15 URLs 437

16 Mechanical Impedances and Admittances for Translatory and Rotatory Systems 443

17 Details about Gyrators and Transformers 445

18 Impedance Values of Grasps 449

References 451

Index 467

Symbols

Symbol	Description	Unit
a	acceleration	$\frac{\mathrm{m}}{\mathrm{s}^2}$
a	stimulus-constant	
a	displacement (Denavit-Hartenberg-parameter)	m
$\mathbf{a}$	vector, summarizing actuator displacement and angles a_i (chapter 8)	
$\delta\mathbf{a}$	virtual displacement of drives (chapter.8)	m
A	area	m^2
A_{Coil}	cross section of a winding	m^2
$A_{Conductor}$	surface of a coiling	m^2
A_{Core}	area of the conductive copper area	m^2
A_{Iron}	cross-section of the iron core	m^2
A_G	cross-section of the airgap of an electromagnetic actuator	m^2
A_n	cross-section of the element n of a magnetic circuit	m^2
$\Delta A_j, j=1,2,3$	surface element	m^2
$A(\omega)$	amplitude response (chapter 7)	dB
$\mathbf{A}$	matrix of a linear system of equations	
α	roll-angle (rotation around x-axis)	degree
α_c	border-angle (chapter 8)	degree
α_i	rotation (Denavit-Hartenberg-parameter) (chapter 8)	degree
α_{VK}	coefficient of thermal expansion	K^{-1}
b_0	transfer factor for differential movements	

Symbol	Description	Unit
B_{Iron}	magnetic flux density in the iron core	$\mathrm{T} = \frac{\mathrm{V \cdot s}}{\mathrm{m^2}}$
B_G	magnetic flux density of an electrodynamic actuator	T
B_n	magnetic flux density in the element n of a magnetic circuit	T
B_r	remanence flux density	T
$\mathbf{B}$	magnetic flux density	T
$\mathbf{B}$	matrix of a linear system of equations	
BH_{max}	energy density	$\frac{\mathrm{J}}{\mathrm{m^3}}$
β	pitch angle (rotation around the y-axis)	degree
c	perceptional constant	
c_{ijlm}	elastic constants	$\frac{\mathrm{m^2}}{\mathrm{N}}$
C, C_Q	capacity	$\mathrm{F} = \frac{\mathrm{A \cdot s}}{\mathrm{V}}$
C_0	photoelastic coefficient of a Bragg-lattice	
C_b	coupling capacity (at mechanical full-stop)	F
C_d	field control constant of an ERF	$\frac{\mathrm{N}}{\mathrm{V^2}}$
C_s	material-specific field constant of an ERF	$\frac{\mathrm{N}}{\mathrm{V^2}}$
$\mathbf{C}$	matrix of a linear system of equations	
$\frac{\Delta C}{C_0}$	capacity change	
$circ$	circumference of a conductive coil	m
d	damping/friction	$\frac{\mathrm{N}}{\mathrm{m \cdot s}}$
$d_{ij,k}$	piezoelectric charge constant	$\frac{\mathrm{V}}{\mathrm{m}}$
d_i	displacement (Denavit-Hartenberg-parameter)	m
$\mathbf{D}$	dielectric displacement/ electrical displacement density	$\frac{\mathrm{C}}{\mathrm{m^2}}$
$\mathbf{D}$	(transmission-)matrix of a linear system of equations	
DL	difference limen/ amplitude resolution	
DOF	degrees-of-freedom	
ΔD	position-discrete resolution	
δ	phase difference (chapter 10)	
e	piezoelectric voltage coefficient	$\frac{\mathrm{A \cdot s}}{\mathrm{m^2}}$
$\mathbf{e}_i$	directional unit vector	
E	E-modulus	$\frac{\mathrm{N}}{\mathrm{m^2}}$
E	electrical field strength (with dielectric material)	$\frac{\mathrm{V}}{\mathrm{m}}$

Symbol	Description	Unit
E_0	electrical field strength (without dielectric material)	$\frac{\mathrm{V}}{\mathrm{m}}$
E_{ref}	reference field strength, with C_s of an ERF being given	$\frac{\mathrm{V}}{\mathrm{m}}$
$\mathbf{E}$	electrical field	$\frac{\mathrm{V}}{\mathrm{m}}$
ε	permittivity ($\varepsilon = \varepsilon_0 \cdot \varepsilon_r$)	$\frac{\mathrm{A \cdot s}}{\mathrm{V \cdot m}}$
ε	relative dielectric constant at constant mechanical tension	$\frac{\mathrm{A \cdot s}}{\mathrm{V \cdot m}}$
ε	remaining error (chapter 7)	
ε_0	electrical field constant ($\varepsilon_0 = 8,854 \cdot 10^{-12} \frac{\mathrm{C}}{\mathrm{V \cdot m}}$)	$\frac{\mathrm{C}}{\mathrm{V \cdot m}}$
ε_r	relative permittivity ($\varepsilon_r = \frac{E_0}{E}$)	
η	basic-viscosity	$\frac{\mathrm{m}^2}{\mathrm{s}}$
f	frequency	Hz
f_0, f_R	resonance-frequency	Hz
f_b	border-frequency	Hz
f_{tot}	sum of all joint degress-of-freedom of a mechanism	
$f_i, i \ldots g$	degree-of-freedom of the ith joint in a mechanism	
f_{id}	sum of all identical bindings in a mechanism	
f_{ink}	dynamics of the detection of all increments for positioning measurement	Hz
$f(.)$	static non-linearity	
F	bearing-/movement-DOF of a mechanism	
F_a	force in direction of a border a	N
F_A	force acting upon the point A	N
F_b	force in direction of a border b	N
F_e	unit force on the mass m of a system	N
F_F	counter-force	N
F_{in}	input-force	N
F_l	force in direction of l	N
F_N	normative force	N
F_{norm}	normal force	N
F_w	transformer input force	N
F_x	force in x-direction	N
F_y	force in y-direction	N

Symbol	Description	Unit
$F_{z,el}$	force in an electrical field in z-direction	N
F_η	velocity dependent viscosity force	N
F_τ	field dependent strain term	N
F_ξ	force at position ξ	N
$\underline{F}$	force	N
$\underline{F}_D$	device related force	N
$\underline{F}_m$	force on mass m	N
$\underline{F}_{noise}$	disturbing force	N
$\underline{F}_{out}$	output/generated force	N
$\mathbf{F}$	force	N
$\mathbf{F}_a$	vector of all force/torques on the driven joints	N
$\mathbf{F}_F$	counter-force	N
$\mathbf{F}_{Lorentz}$	LORENTZ-force	N
$\mathbf{F}_x$	vector of all forces/torques on the tool-center-point	N
$\mathbf{F}_x$	force for displacement in x-direction	N
$\mathbf{F}_y$	force for displacement in y-direction	N
$\mathbf{F}_{z,el}$	force in the electrical field in z-direction	N
ΔF	force-resolution	N
$\Delta F_i, i = 1,2,3$	replacement-forces	N
ϕ	magnetic flux	Wb=V $\cdot$ s
$\phi(\omega)$	phase plot	degree
$\boldsymbol{\Phi}$	phase-angle	rad
$\boldsymbol{\Phi}$	phase-displacement	rad
$\boldsymbol{\Phi}$	stimulus	
g	number of joints in a mechanism (chapter 2)	
g	piezoelectric constant	$\frac{\mathrm{V \cdot m}}{\mathrm{N}}$
$\underline{G}$	transfer function	
$\underline{G}_{CD}$	transfer function of a controller	
$\underline{G}_{Dn}, n \in \aleph$	device-related transfer function	
$\underline{G}_{ED}$	transfer function of a driver (transformation of a force signal in energy)	
$\underline{G}_{FF}$	feedforward-transfer function	
$\underline{G}_{FIP}$	transfer function for the transformation of perception of mechanical oscillations	

Symbol	Description	Unit
$\underline{G}_{FSense}$	transfer function of a force sensor	
$\underline{G}_{Hn}, n \in \aleph$	user-related transfer function	
$\underline{G}_K$	transfer function of a compensator	
$\underline{G}_M$	transfer function of a measurement unit	
$\underline{G}_R$	transfer function of a controller	
$\underline{G}(s)$	transfer function in LAPLACE domain	
$\underline{G}_S$	transfer function of the control/the actuator	
$\underline{G}_{SZ}$	noise transfer function	
γ	yaw angle (rotation around the z-axis)	degrees
$\dot{\gamma}$	shear-rate	s^{-1}
h	height	m
H	magnetic field strength	$\frac{\mathrm{A}}{\mathrm{m}}$
H_c	coercitive field strength	$\frac{\mathrm{A}}{\mathrm{m}}$
H_{Fe}	magnet. field strength of a coil with iron core	$\frac{\mathrm{A}}{\mathrm{m}}$
H_G	magnet. field strength in the air gap of a iron core	$\frac{\mathrm{A}}{\mathrm{m}}$
H_n	magnet. field strength in the element n of a magnetic circuit	$\frac{\mathrm{A}}{\mathrm{m}}$
$\underline{i}$	current	A
i_L	current through and inductance L	A
i_{Source}	source current	A
i_w	transform input current	A
I	moment of inertia	m^4
I_b	current at the input of an operational amplifier	A
$\mathbf{j}$	current density	$\frac{\mathrm{A}}{\mathrm{m}^2}$
$\mathbf{J} = \frac{\partial \mathbf{x}}{\partial \mathbf{a}}$	Jacobi-matrix with actuator DOF $\mathbf{a}$ and DOFs of the tool center points $\mathbf{x}$	
JND	Just-Noticeable-Difference	
k	geometrical design dependent constant of ERFs	$\mathrm{m} \cdot \mathrm{s}$
k	fill-factor of a coil (≥ 1)	
k	gradient in the POPOV-criterion	
k	coupling-factor or k-factor	
k	chain-number in a mechanism	
k	stiffness (mechanical)	$\frac{\mathrm{N}}{\mathrm{m}}$
k_L	stiffness in idle mode	$\frac{\mathrm{N}}{\mathrm{m}}$

Symbol	Description	Unit
K'	user reaction	
K_{krit}	critical amplification	
K_R	controller amplification	
$\tilde{K}$	perceptional space	
ΔK	time dependency of perception curve	dB
$\underline{K}$	force-perception	
$\mathbf{K}$	amplification matrix	
l	length	m
$l_{Conductor}$	length of a conductor	m
l_{Fe}	length of a coil with iron core/of a magnetic circuit	m
l_{Mag}	length of a rare-earth magnet	m
l_n	length of the element n of the magnetic circuit	m
L	inductivity	$\mathrm{H} = \frac{\mathrm{V \cdot s}}{\mathrm{A}}$
LM	Successiveness Limen	s
$\mathbf{L}$	amplification matrix	
λ	mechanical control ratio	
λ	wavelength	m
λ_0	original wavelength at a defined position in the spectrum	m
λ_b	Bragg-wavelength	m
Λ	Bragg-lattice period	m
m	mass	kg
M_α	Input torque in axial direction of the drive	Nm
M_b	bending moment	Nm
$\underline{M}$	torque	Nm
$\underline{M_0}$	torque-source	
μ	movability of a charge-carrier	$\frac{\mathrm{m}^2}{\mathrm{V \cdot s}}$
μ	frictional coefficient	
μ	mean value	
μ	permeability ($\mu = \mu_0 \cdot \mu_r$)	$\frac{\mathrm{V \cdot s}}{\mathrm{A \cdot m}}$
μ_0	magnetic field constant	$\frac{\mathrm{H}}{\mathrm{m}}$
μ_r	relative permeability	
n, N	number $\in \aleph$	
n	elasticity	$\frac{\mathrm{N}}{\mathrm{m}}$

Symbol	Description	Unit
n_0, n_i	refraction index	
n_K	elasticity of the coupling of a mechanical short-circuit	$\frac{\mathrm{N}}{\mathrm{m}}$
$\delta\bar{n}_{effektiv}$	medium value of index modulation	
Δn_{RK}	elasticity of the rotary coupling of a mechanical short-circuit	$\frac{\mathrm{N}}{\mathrm{m}}$
$N_{Conductor}$	number of coilings	
NA	numerical aperture	
$\underline{\omega}$	rotational velocity	s^{-1}
$\underline{\omega}, \underline{\Omega}$	angular velocity	$\frac{\mathrm{rad}}{\mathrm{s}}$
p, P	pressure	$\frac{\mathrm{N}}{\mathrm{m}^2}$
P	polarization	$\frac{\mathrm{C}}{\mathrm{m}^2}$
P_{el}	electrical power loss	$\mathrm{W} = \mathrm{V} \cdot \mathrm{A}$
p_{el}, P_{el}	electrostatic pressure	$\frac{\mathrm{A \cdot V \cdot s}}{\mathrm{m}^3}$
P_{Loss}	power loss	W
P_{mech}	mechanical power	W
P_{Source}	source power	W
P_η	presssure loss due to viscosity	$\frac{\mathrm{N}}{\mathrm{m}^2}$
P_τ	field dependent pressure loss	$\frac{\mathrm{N}}{\mathrm{m}^2}$
π	piezoresistive coefficient	$\frac{\mathrm{m}^2}{\mathrm{N}}$
π_l	piezoresistive coefficient in longitudinal direction	$\frac{\mathrm{m}^2}{\mathrm{N}}$
π_q	piezoresistive coefficient in transversal direction	$\frac{\mathrm{m}^2}{\mathrm{N}}$
Ψ	perception amplitude	
q, Q	electrical charge	$\mathrm{C} = \mathrm{A} \cdot \mathrm{s}$
$q_i, i \in \aleph$	driven joint i	
$\mathbf{q}$	vector of a driven joint q_i	
r	distance	m
r	radius	m
$r_i, i \in \aleph$	active resistors	$\Omega = \frac{\mathrm{V}}{\mathrm{A}}$
R	electrical resistance	Ω
R_0	electrical basic resistance	Ω
R_{coil}	coiling resistance	Ω
$R_i, i \in \aleph$	reference resistance	Ω
R_L	conductor resistance	Ω

Symbol	Description	Unit
R_m	magnetic resistance/reluctance	$\frac{\mathrm{A}}{\mathrm{V \cdot s}}$
R_{Mag}	magnetic resistance of a rare-earth magnet	$\frac{\mathrm{A}}{\mathrm{V \cdot s}}$
R_{mFe}	magnetic resistance of an iron circuit	$\frac{\mathrm{A}}{\mathrm{V \cdot s}}$
R_{mG}	magnetic resistance of an air gap within a magnetic circuit	$\frac{\mathrm{A}}{\mathrm{V \cdot s}}$
R_{Sense}	sensing resistor of a PWM	Ω
$R_{spezf.}$	specific resistance	Ω
$\frac{dR}{R_0}$	relative resistance change	
ΔR_{inch}	position resolution given in dots-per-inch	DPI
ΔR_{mm}	position resolution given in millimeter	mm
ρ	small number ≥ 0	
ρ	density	$\frac{\mathrm{kg}}{\mathrm{m}^3}$
ρ	specific resistance/conductivity	$\Omega \cdot \mathrm{m}$
s	elasticity coefficient at a constant field strength	$\frac{\mathrm{m}^2}{\mathrm{N}}$
s	measure for the level of modulation in an optical fibre	
s	sum of the passive links of a mechanism	
S	external supply rate	
S	mechanical tension/deformation	
S	constraints in a mechanism	
$S_i, i \in \aleph$	ith switch	
S_x	transversal-tension	m
S_z	longitudinal compression	m
σ	conductivity	$\frac{\mathrm{S}}{\mathrm{m}} = \frac{\mathrm{A}}{\mathrm{V \cdot m}}$
σ	standard deviation	
t	time/point in time	s
tr	transmition ratio of a gear	
T	mechanical tension	$\frac{\mathrm{N}}{\mathrm{m}^2}$
T	time constant	s
T_{an}	rise time (100% of the target value are reached for the first time)	s
T_D	rate time	s
T_l	mechanical tension in longitudinal direction	$\frac{\mathrm{N}}{\mathrm{m}^2}$
T_{max}	point in time for $x_{d,max}$	s
T_N	reset time constant	s

Symbol	Description	Unit
T_q	mechanical tension in transversal direction	$\frac{\mathrm{N}}{\mathrm{m}^2}$
T_t	dead time	s
$\underline{T}$	transparency	
$\mathbf{T}$	tension	$\frac{\mathrm{N}}{\mathrm{m}^2}$
$\mathbf{T}$	special case of the homogeneous transformation matrix	
T_ε	transient time until ε	s
TCP	tool center point	
τ	shear force	N
τ	time constant of the step response of an electrical transmission system ($\tau = \frac{L}{R}$)	s
τ_{ERF}	yield stress of an electrorehological fluid	$\frac{\mathrm{N}}{\mathrm{m}^2}$
$\tau_{F,d}$	dynamic yielding point	$\frac{\mathrm{N}}{\mathrm{m}^2}$
$\tau_{F,s}$	static yielding point	$\frac{\mathrm{N}}{\mathrm{m}^2}$
τ_{MRF}	yield stress of a magnetorheological fluid	$\frac{\mathrm{N}}{\mathrm{m}^2}$
τ_y	yield stress	$\frac{\mathrm{N}}{\mathrm{m}^2}$
θ	rotational angle (Denavit-Hartenberg-parameter)	degree
Θ	magnetomotive force	A
Θ	mass inertia	$\mathrm{kg} \cdot \mathrm{m}^2$
Θ_c	accepted angle	degree
ϑ	temperature	K
$\underline{u}$	voltage	V
u_C	voltage at a capacitor	V
$u_{Ri}, i \in \aleph$	voltage at a resistor R_i	V
u_{Source}	source voltage	V
$\mathbf{u}$	multidimensional input value of a linear system	
U	electrical voltage	V
U_b	supply voltage	V
U_{in}	input voltage	V
U_{ind}	induced voltage	V
$U_{pull-in}$	voltage at which the pull-in happens	V
U_{Sense}	current-proportional voltage at R_{Sense}	V
v_a	velocity at point A	$\frac{\mathrm{m}}{\mathrm{s}}$
v_e	input velocity of a system	$\frac{\mathrm{m}}{\mathrm{s}}$

Symbol	Description	Unit
v_{explor}	velocity of a movement	$\frac{\mathrm{m}}{\mathrm{s}}$
v_{in}	input velocity	$\frac{\mathrm{m}}{\mathrm{s}}$
v_m	velocity of a mass m	$\frac{\mathrm{m}}{\mathrm{s}}$
v_{max}	maximum velocity	$\frac{\mathrm{m}}{\mathrm{s}}$
$\underline{v}$	velocity	$\frac{\mathrm{m}}{\mathrm{s}}$
$\underline{v}_0$	velocity of a linear movement	$\frac{\mathrm{m}}{\mathrm{s}}$
$\underline{v}_D$	movement velocity of a device	$\frac{\mathrm{m}}{\mathrm{s}}$
$\underline{v}_H$	user-related velocity	$\frac{\mathrm{m}}{\mathrm{s}}$
$\underline{v}_{ind}$	velocity of an induced movement	$\frac{\mathrm{m}}{\mathrm{s}}$
$\underline{v}_{out}$	ouput velocity/ generated velocity	$\frac{\mathrm{m}}{\mathrm{s}}$
$\underline{v}_{spo}$	velocity of a spontanious movement	$\frac{\mathrm{m}}{\mathrm{s}}$
$\underline{v}_Z$	velocity at a mechanical impedance	$\frac{\mathrm{m}}{\mathrm{s}}$
V	magnetic tension	A
V	volume	m^3
V_{ERF}	volume of a electrorheological fluid	m^3
V_{MRF}	volume of a magnetorheological fluid	m^3
$V(\mathbf{x})$	scalar nonlinear positive definite stoarge function of system states $\mathbf{x}$	
$\mathbf{V}$	pre-filter matrix	
$\dot{V}$	volume flow	$\frac{\mathrm{m}^3}{\mathrm{s}}$
ΔV	volume-element	m^3
w	general value for in- and ouput values	
w	wave-amplitude of the stator	m
W_{el}	electrical work/energy	$\mathrm{J} = \frac{\mathrm{kg}\cdot\mathrm{m}^2}{\mathrm{s}^2}$
$W_{el,ERF}$	electrical energy of an electrorheological fluid	J
$W_{el,MRF}$	electrical energy of a magnetorheological fluid	J
W_{mag}	magnetic energy	J
W_{mech}	mechanical energy	J
δW_α	virtual work of a propulsion cylinder	J
δW_x	virtual work of an actuator	J
x	distance	m
$x_d(t)$	control deviation of a closed-loop control circuit	
$x_{d,max}$	maximum overshoot of a closed-loop control circuit	

Symbol	Description	Unit
$\underline{x}$	displacement/translation	m
$\mathbf{x} = (x, y, z)$	cartesian coordinates	
$\mathbf{x}$	inner states of a linear system	
$\mathbf{x}$	vector summarizing translations and angles x_j of the control handle	
$\delta\mathbf{x}$	virtual displacement of the tool-center-point	m
Δx	position resolution	m
$\underline{\Delta x}$	position of a point in three-dimensional space	
X	transformation constant	
ξ	displacement	m
ξ_G	air-gap length in a magnetic circuit	m
$\underline{d\xi}$	spatial displacement	
y	control value	
$\mathbf{y}$	multidimensional output value of a linear system	
Y	gyratoric transformation constant	
$\underline{Y}$	admittance	$\frac{\text{m}}{\text{N·s}}$
$\underline{Y}_H$	user-related admittance	$\frac{\text{m}}{\text{N·s}}$
z_i	disturbance variable	
$z(t)$	disturbance variable in a closed-loop control circuit	
$\underline{Z}$	impedance (usual mechanic)	$\frac{\text{N·s}}{\text{m}}$
$\underline{Z}_D$	display-/interface-impedance	$\frac{\text{N·s}}{\text{m}}$
$\underline{Z}_D$	impedance of the device	$\frac{\text{N·s}}{\text{m}}$
$\underline{Z}_H$	user-impedance	$\frac{\text{N·s}}{\text{m}}$
$\underline{Z}_{in}$	impedance as input value	$\frac{\text{N·s}}{\text{m}}$
$\underline{Z}_{max}$	maximum impedance/ full load	$\frac{\text{N·s}}{\text{m}}$
$\underline{Z}_{min}$	minimum impedance in idle movement	$\frac{\text{N·s}}{\text{m}}$
$\underline{Z}_{out}$	impedance as output value	$\frac{\text{N·s}}{\text{m}}$
$\underline{Z}_{rot}$	rotatory impedance of a motor ($= \frac{M}{\underline{\alpha'}}$)	$\frac{\text{N·s}}{\text{m}}$
$\underline{Z}_{transl}$	translatory impedance of a motor ($= \frac{F}{\underline{v}}$)	$\frac{\text{N·s}}{\text{m}}$
$\underline{Z}_{width}$	impedance width, measure of performance ($\underline{Z}_{width} = \underline{Z}_{max} - \underline{Z}_{min}$)	$\frac{\text{N·s}}{\text{m}}$

Structure of this Book

This book subdivides into two major parts. **Part I** "Some Basics of Haptics" presents the sense of touch- and movement from different scientific and popular perspectives. It defines the specific area, in which context "haptics" will be used throughout this book. Starting from the definition of terms, an introduction into the biological basics of haptic perception is given, to gain some insight into the specialties of this spatially distributed sensing organ, and sharpens the reader's own senses for generating technical illusions for haptics. Afterward some basic structures and classes of haptic systems are introduced, and first requirements on the technical design are derived. The first part ends with a discussion about the possibilities to quantify haptic perception illustrated with several examples.

Part II "Designing Haptic Systems" deals with the actual technical design process in all facets relevant for haptic systems. Following a chapter on requirement identification, methods for control engineering support the identification of interdependencies between the components within a system. A chapter on the engineering of kinematics, with an emphasis on the frequently - in the context of haptic devices - used parallel kinematics, and their calculation builds the basis for the next steps in the design process. The most compelling chapter of part II covers actuator design. All important actuation principles are introduced and discussed for their application in haptic devices. Common actuation principles are detailed and described up to a point, where any engineer should be able to start with own designs. Within the chapter about force-sensors technology especially relevant for telemanipulation and closed-loop haptic devices is discussed. The next chapter adds information about positioning sensors and realizable resolutions. A presentation of relevant interfaces closes the hardware design part. The subsequent chapter about software design introduces terminology from a neighboring discipline, which is very relevant for any device engineer - especially in the context of simulator and virtual-reality environments. The part ends with some finalizing remarks about the design of haptic systems.

Part I
Some Basics of Haptics

"*This page left intentionally blank.*"

The design of haptic devices aims at deceiving the human sense of movement and touch. In the first part of the book, a general view of this task, starting with the terminology and biology of haptics is given. Additionally, the application of psychophysical insights into technological problems is discussed.

- Chapter 1: The **motivation** for the design of haptic systems is presented. The significance of haptics in a social context and in professional life is given in order to emphasize the developer's responsibility when fooling this sense.
- Chapter 2: A specific **terminology** is introduced and the understanding of various disciplines participating in haptic science and research is detailed.
- Chapter 3: A basic understanding of the **biological receptors** responsible for haptic perception is conveyed. This enables the reader to assess the parameters of the technological influence and enhances his sensibility for the complexity of the haptic sense.
- Chapter 4: A technically motivated **model** of the human user as mechanical load - as a consequence of the technological influence on haptic perception - is given (section 4.2) for the system, and a frequency dependent parameter (section 5) for its perception. Together this results in a model of the human-machine interaction and the perception of dynamic mechanical, physical magnitudes.
- Chapter 5: The most important **classes** of haptic devices based on this model are introduced.

In the second part of the book the focus switches from the haptic sense to the technical design of haptic devices.

"*This page left intentionally blank.*"

Chapter 1
Motivation and Application of Haptic Systems

THORSTEN A. KERN

1.1 The Meaning of Haptics from a Philosophical and Social Viewpoint

Haptics describes the sense of touch and movement. An engineer tends to describe haptics in terms of forces, elongations, frequencies, mechanical tensions and shear-forces. This of course makes sense and is important for the technical design process. However haptics is more than that. Haptic perceptions range from minor interactions in everyday life, e.g., drinking from a glass or writing this text, to a means of social communication, e.g. shaking hands or giving someone a pat on the shoulder, and very personal and private interpersonal experiences. This chapter deals with the spectrum and influence haptics has on the human being beyond technological descriptions. It is also a hint for the development engineer, to be responsible and conscious when considering the capabilities to fool the haptic sense.

1.1.1 Haptics as a Physical Being's Boundary

Haptics is derived from the Greek term "haptios" and describes "something which can be touched". In fact the consciousness about and understanding of the haptic sense has changed many times in the history of humanity. ARISTOTELES puts the sense of touch in the last place when naming the five senses:

1. sight
2. hearing
3. smell
4. taste
5. touch

T.A. Kern (ed.), *Engineering Haptic Devices*, DOI 10.1007/978-3-540-88248-0_1,
© Springer-Verlag Berlin Heidelberg 2009

Nevertheless he attests this sense a high importance concerning its indispensability [3]:

The social estimation of the sense of touch experienced all imaginable phases. Frequently it was afflicted with the blemish of squalor, as lust is transmitted by it:

"*Sight differs from touch by its virginity, such as hearing differs from smell and taste: and in the same way their lust-sensation differs* [289]."

It was called the sense of excess [78] . In a general subdivision between lower and higher senses, touch was almost constantly ranged within the lower class. In western civilization the church once stigmatized this sense as forbidden due to the pleasure which can be gained by it. However in the 18th century the public opinion changed and KANT is cited with the following statement [126]:

"This sense is the only one with an immediate exterior perception; due to this it is the most important and the most teaching one, but also the roughest. Without this sensing organ we would be able to grasp our physical shape, whose perception the other two first class senses (sight and hearing) have to be refered to, to generate some knowledge from experience."

KANT thus emphasizes the central function of the sense of touch. It is capable of teaching the spatial perception of our environment. Only touch enables us to feel and classify impressions collected with the help of other senses, put them into context and understand spatial concepts. Although stereoscopic vision and hearing develop early, the first-time interpretation of what we see and hear, requires the connection between both impressions perceived independently and information about distances between objects. This can only be provided by a sense, which can bridge the space between a being and an object. Such a sense is the sense of touch. The skin, being a part of this sense, covers a human's complete surface and defines his or her physical boundary, the physical being.

Wearing glasses is another fascinating example of the effect of the relationship between distance and perception. Short- sightedness requires glasses that demagnify the picture of the environment on the retina due to the distance between eyeball and lenses. Shortsighted people have a different view of size, e.g. concerning their own body height, dependent on whether they wear glasses or contact lenses. At every change between both optical aids the perception of their body has to adapt. Dependent on a person's kind of defective vision this is a consciously perceivable process. It can be performed within seconds by using the well known references of one's own arms which touch things or one's legs which walk.

Especially in the 20th century art deals with the sense of touch and plays with its meaning. Drastically the furry-cup (fig. 1.1) makes you aware of the significance of haptic texture for the perception of surfaces and surface structures. Whereas the general form of the cup remains visible and recognizable, the originally plane ceramic surface is covered by fur. “Fighting the mud” (fig. 1.2) remembers you that not only hands and fingers are relevant for haptic perception, but that the whole body surface is able to touch and feel. In 1968 the “Pad- and Touch-Cinema” (fig. 1.3) allowed visitors to touch VALIE EXPORT’s naked skin for 12 seconds through a box being covered by a curtain all the time. According to the artist this was the only valid approach to experience sexuality without the aspect of voyeurism [70]. These are just a few examples of how art and artists played with the various aspects of haptic perception.



Fig. 1.1 MERET OPPENHEIM: furry-cup 1936 [70, 186].

Also with virtual worlds and surroundings, haptic interaction has characteristics of artistry. This is repeatedly demonstrated by expositions during the Worldhaptics Conferences. At the same time, Prof. ISHII from MIT Media Laboratory or the Graduate School of Systems and the Information Engineering group of the University of Tsukuba of Prof. IWATA demonstrate startling exhibits (fig. 1.4) of “tangible user interfaces” (TUI). These interfaces couple visual displays with haptically reconfigurable objects to provide intuitive human-machine interfaces. There is much more to find when the senses are sharpened to search for it.

The sense of touch can be a lot of things, e.g. a limitation of the physical being, which helps to assess distances and calibrate other senses like vision, as well as a means of social communication and a mediator of very personal experiences .Additionally it is - like all the other senses - a target of art which makes us aware of the importance of haptic experiences by fooling, distorting and emphasizing them. . Besides these facets of the haptic sense, its function and its dynamic mechanical



Fig. 1.2 KAZUO SHIRAGA: Doro ni idomu (Fighting the mud) 1955 [70][217].



Fig. 1.3 VALIE EXPORT: Pad- and Touch-Cinema 1968 [70].

properties are also very impressive. Haptic perception in all its aspects is presented in the following section.

1.1.2 Formation of the Sense of Touch

As shown in the prior section, the sense of touch has numerous functions. The knowledge of these functions enables the engineer to formulate demands on the technical system. It is helpful to consider the whole range of purposes the haptic sense serves. However, at this point we do not yet choose an approach by measuring its characteristics, but observe the properties of objects discriminated by it.



Fig. 1.4 Example for "Tangible Bits", with different data streams accessible by opening bottles. In this case single instrumental voices are combined to a trio [109] .

The sense of touch is not only specialized on the perception of the physical boundaries of the body, as said before, but also on the analysis of surface properties. Human beings and their predecessors had to be able to discriminate e.g. the structure of fruits and leaves by touch, in order to identify their ripeness or whether they were eatable or not, like e.g. a furry berry among smooth ones. The haptic sense enables us to identify a potentially harming structure, like e.g. a spiny seed, and to be careful when touching it,in order to obtain its content despite its dangerous needles. For this reason, the sense of touch has been optimized for the perception and discrimination of surface properties like e.g. roughness. Surface properties may range from smooth ceramic like or lacquered surfaces with structural widths in the area of low μm, to somewhat structured surfaces like coated tables and rough surfaces like coarsely woven cord textiles with mesh apertures in the range of several millimeters. Humans have developed a very typical way how to interact with theses surfaces enabling them to draw conclusions based on the underlying perception mechanism. A human being moves his or her finger along the surface (fig. 1.5), allowing shear forces to be coupled to the skin. The level of the shear forces is dependent on the quality of the frictional coupling between the object surface and the skin. It is a summary of the tangential elasticity of the skin depending on the normal pre-load resulting from the touch F_{norm} and the velocity F_{explr} of the movement and the quality of the coupling factor μ.

Everyone who has ever designed a technical frictional coupling mechanism knows that without additional structures or adhesive materials viscous friction between two surfaces can hardly reach a factor of $\mu_r = 0.1$. Nevertheless nature, in order to be able to couple shearing force more efficiently into the skin, has "in-

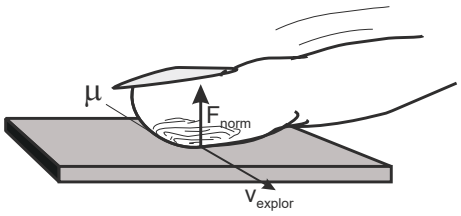


Fig. 1.5 Illustration for the interaction of movements, normal forces on the finger pad and frictional coupling.

vented" a special structure at the most important body-part for touching and exploration: the fingerprint. The epidermal ridges couple shearing forces efficiently to the skin, as by the bars a bending moment is transmitted into its upper layers. Additionally these bars allow form closures within structural widths of similar size, which means nothing else but canting between the object handled and the hand's skin. At first glance this is a surprising function of this structure. When one looks again, it just reminds you of the fact that nature does not introduce any structure without a deeper purpose. Two practical facts result from this knowledge: First of all the understanding of shear-forces' coupling to the skin has come into focus of current research [65] and has resulted in an improvement of the design process of tactile devices. Secondly, this knowledge can be applied to improve the measuring accuracy of commercial force sensors by building ridge-like structures [275]. Additional details of the biological basics of tactile perception are given in chapter 3.

Consequently the sense of touch, as said before, has been developed for the discrimination of surface structures. Although the skin may be our most sensitive organ, it is still not the only haptically relevant one. Additional receptors are located within muscles and joints, which enable us to get an impression of acting forces. Anyone who has ever lifted a four pound weight (e.g. a well filled pitcher) with an outstretched arm in a horizontal position, will have little recollection of the tactile surface properties of the handle. The much more impressive experience of such an experiment is the tensing up of the muscles, their slowly increasing fatigue and the resulting change in the lifting angles of the joints. This is called "kinaesthetic perception". Whereas tactile perception describes forces ($\approx 5\,\mathrm{mN}..5\,\mathrm{N}$) and elongations between skin and object which are low in amplitudes ($\approx 1\,\mu\mathrm{m}..1\,\mathrm{mm}$) and high in frequencies ($\approx 10\,\mathrm{Hz}..1000\,\mathrm{Hz}$), kinaesthetic perception happens within muscles and joints at higher forces but with lower dynamics ($\approx$ static..10 Hz). This enables the human being and every other biological system with a firm supportive structure - may it be bones or shells of chitin - to perform coordinated movements and targeted interactions with its environment. While tactile perception generates similar impressions during passive (e.g. a relative movement between a static finger-tip and a moving surface) and active (e.g. a relative movement between a static surface and a moving finger-tip) movement, kinaesthetic perception is more complex and influ-

enced by additional factors. The human being is able to change deliberately his or her mechanical properties. A handshake of the same person can be firm and rigid, but it may be also loose and amicable. The coupling between muscles, joint position and perception enables us to consciously influence the kinaesthetics of ourselves, and to influence the intensity of our kinaesthetic perception in one and the same situation. This makes us capable of blocking a blow with the same hands we use to rock a baby to sleep. It gives us the ability to touch a structure carefully before we grasp it firmly. The borders between action and reaction, active and passive become blurred in the perspective of kinaesthetics. The awareness of this fact is important for the requirements on systems with closed-loop control, which are important for the design of haptic devices (chapter 5). At the same time this adaptability of the human being and the connected ambiguity of the system's borders are a significant challenge for the design of a technical device.

1.1.3 Special Aspects of the Design Process

The design of any technical system always includes a long chain of compromises. The achievement of the engineer lies in the selection of those compromises which ensure that existing requirements are still fulfilled. Often these compromises are financially motivated - a product should be inexpensive during the production process without losing performance. Concerning these demands, an optimization of systems with interfaces to other purely technical systems is often elegantly possible. The technical systems are quite exactly known as to their characteristics and a technical design can anticipate these characteristics with a certain security margin. Thus the interpretation of a sensor capturing the rotation of a wheel, e.g. a speedometer, is a relatively clear task. The necessary speeds are known, and disturbance variables like temperature areas as well as humidity. can be identified Alternatively they can be measured with high exactness. It is also relatively easy to identify the requirement of measuring a two-dimensional movement of a human operated device on a level surface - e.g. a computer mouse. The temperature range of the appliance is known; the disturbance variables are limited to the optical measurement path and the mechanical surface state and can easily be investigated. Only the speed is not given as precisely as by a technical system. It results from the consideration about the maximum speeds a human hand can reach. Here uncertainties become evident, soon. Although the dynamics of human movement can be measured - technically, a high variance between different people will be observed. This variance also concerns the technical requirements of any object used by humans, and may it only be the physical dimensions of tables and chairs. Dealing with such variances, matching measuring methods and statistical analysis methods have entered anthropometric modeling up to ergonomic design of work-places [153] as well as ergonomic standardization ISO norms 9241/DIN 33 402 [1]. The science of anthropometrics applies

[1] Ergonomics is the science of the conformity of human work with natural law.

to static (lengths, dimensions) and dynamic (speeds) cases. As a matter of fact: Every human's-applicable characteristic value is affected by such a wide variance that with the information of ergonomic or anthropometric data only a proportional estimation can be made. These estimations are called percentiles (fig. 1.6). A percentile is a percentage of the totality of the data subject to analysis (e.g. European female children between 10 and 15 years) and, depending on the context, encloses all people who exceed or are below the percentage.

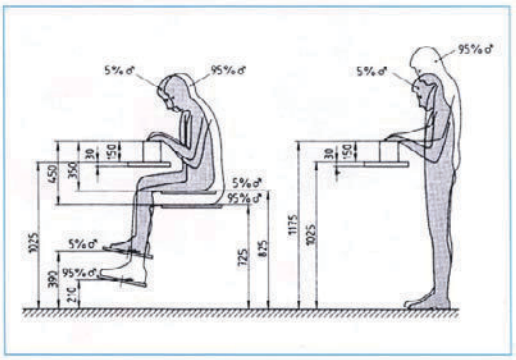


Fig. 1.6 Anthropometric design for sitting and standing work places considering the 5% and the 95% percentile according to DIN 33406.

With regard to the description of body masses and dynamics the description via percentiles introduced above is well established, as it fits quite well the natural variance of people. With regard to the description of senses and their performances average values are more common, e.g. when using a threshold [2]. Thresholds themselves are a key parameter in finding physical values to quantify human performance. Derived from such values the technical system's requirements like amplitude, amplitude change or dynamics can be employed for deceiving a human sense and for generating a "realistic" or "sufficient" haptic impression. The choice of words already shows that requirements seldom comprise a concrete verifiable measurement. They mostly represent a well-known structure, so that a group of people - or just the superior or the board of directors - is content with its haptic impression. For the design engineer this is an unsatisfying benchmark. Alternatives will be discussed to a large extent. in the course of this book and especially in chapter 6.

[2] The use of mean values for the description of the performance of sense organs is doubtful as it would be surprising that our senses variability are met best with a Gaussian distribution. Nevertheless it is found in almost all publications. At least reporting variance should be provided when talking about mean values to improve the quality of the results for interpretation. A concept comparable to the approach of percentiles can hardly be found in the relevant literature. It is likely that the number of experiments and available data is not yet sufficient to use this concept.

1.2 The Significance of Haptics in Everyday Professional Life

The importance of haptics for professional life differs dependent on the profession considered. In handcraft or manual trades the word ‘hand’ already implies the relevance of haptics for performing these jobs. No bricklayer, carpenter, butcher, plumber or barber would be able to do his job, if the sense of touch did not give then important information about the material they work on. May that be the hair they hold between their fingers, the humidity of the wall (as a change of heat transmission), the cable core within the insulation, the difference between tendons and muscles, the graining of pine and beech trees, the consistency of mortar. Even with today’s state-of-the-art technology the involvement of man increases with the required complexity and carefulness of a manual work. With this involvement and the use of sense of touch the tools usually become less complex. Whereas during archaeological excavation a first layer of earth is removed with an excavator, when approaching a hidden structure a shovel will be used or maybe a spatula or for precision work a brush or even the bare hands. However even in handcraft jobs machines of increased flexibility made people turn away from the workpiece and its haptic properties. Today master craftsmen criticize apprentices either for not having anymore a sensation for materials and their properties or for lacking the information -based technological know-how for the control of machines. By optimizing the interface between manual work and machine-programming, engineers try to overcome this gap. But in other areas of professional life, not only in jobs carrying the word “manual” in their name, the loss of the sense of touch for everyday work has already taken place.

1.2.1 The Sense of Touch in Everyday Medical Life

In many medical disciplines high manual skills are required. The capabilities of the sense of touch are necessary for diagnostics and therapy, be it for the identification of skin diseases, the diagnosis of joints, and the palpation of inner organs from the outside or via natural openings; or for a direct surgical application like the transplantation of a heart, the sawing of the cranium or the punctuation of the spinal cord. The sense of touch transmits a plurality of information about texture, elasticity and temperature to the medical professional - information which would either be inaccessible or not so easily accessible in other ways. Nevertheless, in certain situations it is necessary to substitute the sense of touch in diagnosis and therapy. Via magnetic resonance imaging e.g. tendons and menisci of the knee can be visualized. Thus a demanding manual examination of the joints’ movement range is not necessary; especially as performing the procedure and interpreting the haptically felt data requires experience and still leaves room for misinterpretation. Additionally the results of a manual investigation are harder to explain to the patient than the distinctiveness of a real image. However, when comparing the expenses of both

diagnostic procedures, the precedence should be given to the haptic diagnosis. A compromise can be seen in devices like the “Wristalyzer” [77]. This device either puts varying loads on a moving joint - the wrist - or actively moves it, while dynamically measuring the angle vs. displacement curves. Additionally it acquires a complete electro-myography of the muscles. Besides for diagnosis, devices of this kind are already planned for therapy. By actively generating forces and torques, they can be used for the training of all joints of our extremities, of the cervical spine and of the pelvis. Considering all these factors, there seems to be a tendency for the mechanization in diagnostics and therapy. In orthopedic areas there is, however, still some room to discuss its necessity, whereas in surgery there is an urgent need for mechanization which, however, leads to a loss of haptic impressions. After surgical interventions like e.g. an appendectomy, the wish for small wounds and scars for medical and cosmetic reasons has therefore led to the design of laparoscopic instruments (fig. 1.7). Simply by their length, mass and stiffness they also resemble a filter for the haptic information. This decoupling between patient and surgeon has found its temporary climax in the DaVinci system (fig. 1.8) - a laparoscopic telemanipulation system without force feedback. This loss of the sense of touch during surgical (or any other internistic) interventions is obvious and regrettable. As a result numerous research projects were and still are focusing on an adequate substitute for the direct haptic interaction by alternative technologies [73] or improved instruments with integrated force-feedback [209] (fig. 1.9).



Fig. 1.7 Rigid laparoscopic instrument by Karl Storz.

1.2.2 The Sense of Touch in the Cockpit

Besides the aim of getting information which is already mechanically available (elasticity, surface structure, etc.), there is the necessity to provide artificially generated tactile data in addition to overloaded visual or auditory senses in information-

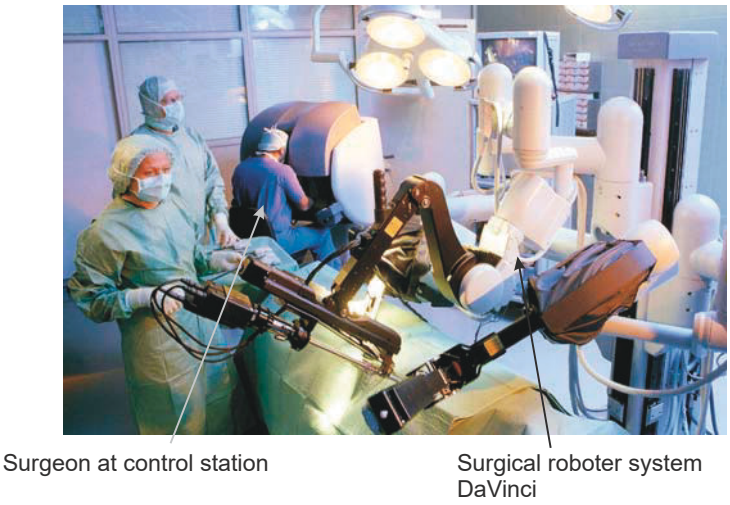


Fig. 1.8 Surgical telemanipulator DaVinci® by Intuitive Surgical, installation in Munich.

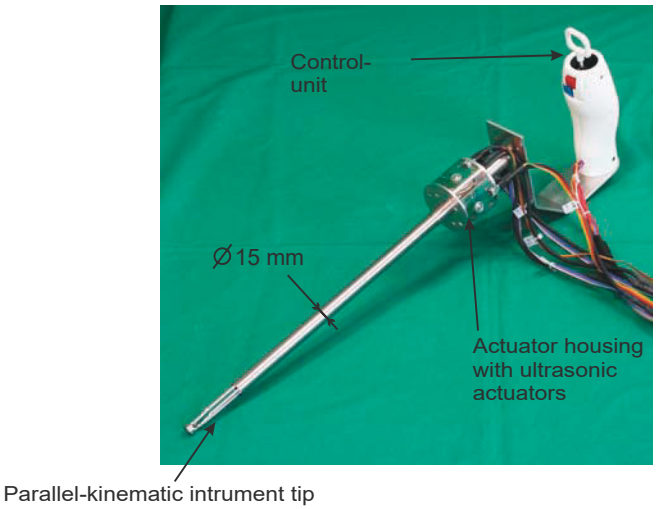


Fig. 1.9 Functional muster of a hand-held laparoscopic telemanipulator with increased number of degrees of freedom at the instrument's tip, such as a prepared intracorporal force measurement with haptic feedback on the control unit [209].

loaded working places. Such working places can be found in control stations where the human has to make time critical and responsible decisions, e.g. within a jet, airplane or at the steering wheel of a common car. The designers of a cockpit typically choose between visual, acoustic and haptic transmission paths. Even the choice of a scroll-wheel with hard stops instead of a pure incremental sensor is influenced by the knowledge that a selection within a certain range can be much faster done if the limits of this range are explicitly given [14]. Control knobs like the i-drive in a BMW allow a reconfiguration of its haptic properties during operation. Warning signals are already given via vibrating motors or so called “tactons”. Especially in the military area a complex spatial orientation based on vibrating clothing (fig. 1.10) for marines and flight-personnel is subject to actual research [267, 115], whereas active sidesticks in military and civil airplanes and vibrating braking assistance or in-lane guidance in cars are already established.



Fig. 1.10 West equipped with vibrators for the spatial coding of positioning and bearing data (TNO, Netherlands) [267].

1.2.3 The Sense of Touch at the Desk

There is hardly any other job where the sense of touch has lost so much of its significance than in the office. Just a few decades ago the use of paper, pens in a large variety, rulers, folders and files was a joyful source of haptic information for the sense of touch. Today the haptic interface to an office working place is defined by a keyboard and a mouse. Due to this extreme focus on a single type of haptic interface for a variety of things, the ergonomics of a keyboard is of extraordinarily high importance. Besides the switching characteristics of the key itself, the surface structure and the tactile markers on the letters F and J (fig. 1.11) and the size of the key are necessary and considerable design criteria. ISO 9241-400 defines clear

decision paths for both, the designer and the buyer of keyboards. Nevertheless it is beyond doubt that major ergonomic improvements are not done by the optimization of keyboard and mouse, but by improvements of office software ergonomics. Contrary to many cases where the term "interface" refers only to the graphical interface, RASKIN's "The Humane Interface" [203] is a decided and enjoyable collection of software with unergonomic graphical interfaces offering methods and design criteria for their improvement.

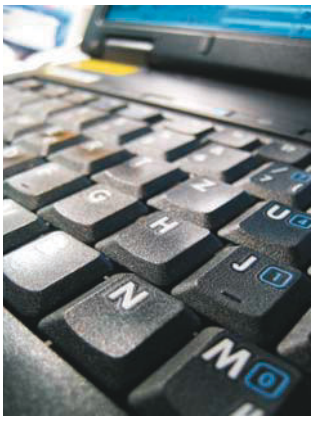


Fig. 1.11 Keyboard of the author with tactile markers on the letters J and F.

1.2.4 The Sense of Touch in Music

If regarded from an abstract standpoint, haptic sense and acoustic perception have multifarious parallels. Both are sensitive to the perception of mechanical oscillations and cover a comparable frequency range. Thereby the haptic sense rather perceives frequencies covering two decades below 1 kHz, whereas the acoustic sense rather perceives frequencies up to two decades above 100 Hz. Music quite often makes use of these parallels which may be used to perceive the oscillations of the string of a valuable violin or harp; or to touch the soft vibration of a wind instrument giving a low A. They are even to be found in studio technology. Devices like the "ButtKicker" (fig. 1.12) from *The Guitammer Company* are electrodynamic actuators which are used as tactile feedback devices during concerts. They transmit the lower frequency range to the drummer giving the rhythm of the band without drowning his own instrument. Additionally the acoustic pressure for the musicians is reduced, as they may not necessarily want to be exposed to the same loudness as their excited audience. These kinds of actuators are also suitable for e.g. the couch in a home

cinema or chairs in front of gaming PCs to increase the perception of bass-intense effects. Here again, the tactile effect is of similar intensity as the perception of a bass impulse, connected with the advantage that little acoustic pressure is emanated resulting in almost no disturbing noise for people around.



Fig. 1.12 Electrodynamic actuator "ButtKicker" for generating low-frequency oscillations on a drum-stool.

Chapter 2
Terminology

THORSTEN A. KERN

In the introduction a number of terminologies originating from the context of haptic science and device design has already been used. In this chapter a systematic introduction into the area of designing haptic devices begins. The following sections explain the scientific and industrial disciplines participating in the research and development of haptic devices. Afterward terms and their definitions are introduced and illustrated by examples how to characterize haptic systems based on some concrete technical devices.

2.1 Scientific Disciplines as Part of Haptic Research

In haptic science there are three groups of interest (fig. 2.1) with quite fluent borders in between: Scientists working within the area of "haptic perception" proceed according to strictly deductive scientific principles: Resulting from an observation a hypothesis is derived. For this hypothesis an experiment is designed by testing the point of the hypothesis by the exclusion of other varying parameters. As a result the hypothesis is veri- or falsified leading to a new and improved hypothesis.

Research in the area of "haptic perception" is done by two scientific disciplines: Psychophysics and Neurobiology. Psychophysics deals with the analysis of the impression of physical stimuli - in the case of haptic perception this mainly refers to oscillations and forces of different spatial orientation. The aim of psychophysics is to create a model explaining perception. Neurobiology observes biologically measurable connections and analyzes the direct conversion of physical stimuli into neuronal signals and their processing within the brain. Both disciplines complement

T.A. Kern (ed.), *Engineering Haptic Devices*, DOI 10.1007/978-3-540-88248-0_2,
© Springer-Verlag Berlin Heidelberg 2009

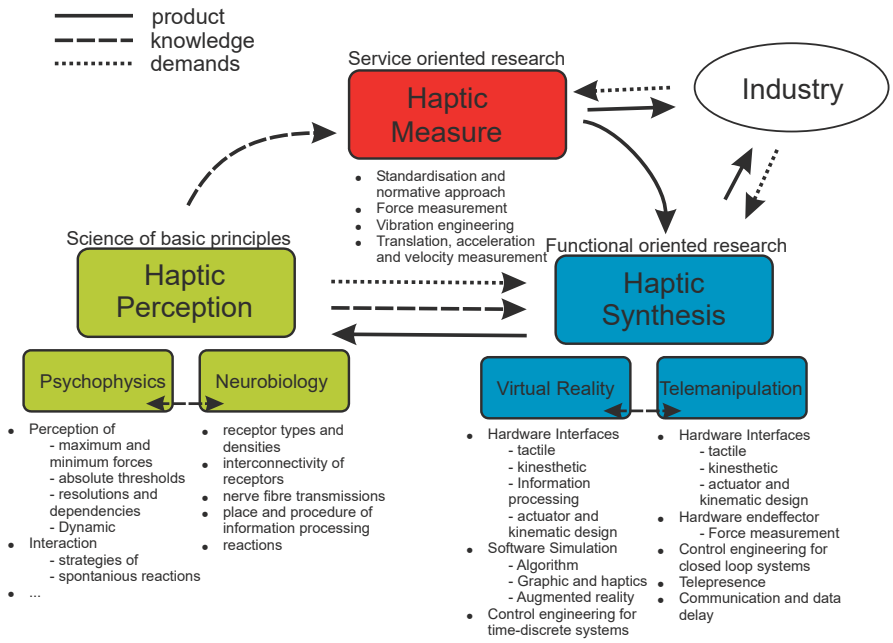


Fig. 2.1 Overview about the disciplines participating in haptic research.

each other so that the neuronal observation should be able to explain a part of the psychophysical model and vice versa. These scientific disciplines formulate technical tasks for the preparation of experiments which are processed by two groups interested in “haptic synthesis” or “haptic measurement”, respectively

On an alternative track both groups get assignments from industry making themselves use of the knowledge gathered by research on haptic perception. These groups work according to engineering solution strategies An assumption of requirements is derived from a technical question based on the current state of knowledge. A functional prototype and later a product to fulfill the requirements is designed in a developmental process accompanied by a continuous tracking of the prior assumptions and their meaning. Then the product obtained can be used for the analysis of psychophysical questions, or, respectively as a a product of the gaming-, automotive or aviation industry.

In the case of the generation of haptic impressions for Virtual-Reality (VR) applications the technical requirements typically ask for tactile, kinaesthetic or combined feedback systems. In that area the emphasis is on the correct choice of actuators, control and driver electronics and on the processing and transmission of signals. Due to the coupling of devices and time-discrete simulation systems a consideration of discretization-effects and their influence on the haptic quality of the impression is necessary. In the case of telemanipulation systems technical challenges are com-

parable. The main difference lies in the necessary measurement technology for the acquisition of haptic object properties. Additionally, the control engineering questions are more complex, as this area typically deals with closed-loop systems with unknown loads on both ends.

2.2 Terms and Terminology Used for the Description of Haptic Systems

The definition of the terminology within the context of haptic systems is subject to the current ISO 9241-910 norm. Many of the definitions used in this book follow the terminology presented there. According to the author's experience, all these terminologies have the status of recommendations shared by a large number of researchers associated with the haptic area. However, there is no binding consent for their usage within the haptic community, so that many actual and future papers differ from the definitions presented above. The nomenclature mentioned here is based on prior publications to this material, especially by HAYWARD [90], COLGATE [176], HANNAFORD [85], BURDEA [34], ADAMS [2] and many papers by other authors.

2.2.1 Basic Concepts of Haptics

Haptics means the combined sensation of mechanical, thermal and noci-perception (fig. 2.2). It is more or less defined by the exclusion of the optical, acoustic, olfactory and gustatory perception from the sum of sensory perceptions. As a result haptics consists of nociceptive, thermoceptive, kinaesthetic and tactile perceptions. The sense of balance takes an exceptional position as it is not counted among the five human senses having receptors of their own Yet, it really exists making use of all other senses' receptors, especially the haptic ones.

Haptics describes the sensory as well as the motor capabilities within the skin, joints, muscles and tendons.

Tactile means the mechanical interaction with the skin. Therefore tactile perception is the sensation of exclusively mechanical interaction. Please note that tactile perception is not exclusively bound to forces or movements.

Kinaesthetics describes both, actuatory and sensory capabilities of muscles and joints. It refers to their forces, torques, movements, positions and angles. As a result any kinaesthetic interaction has a tactile component due to this definition.

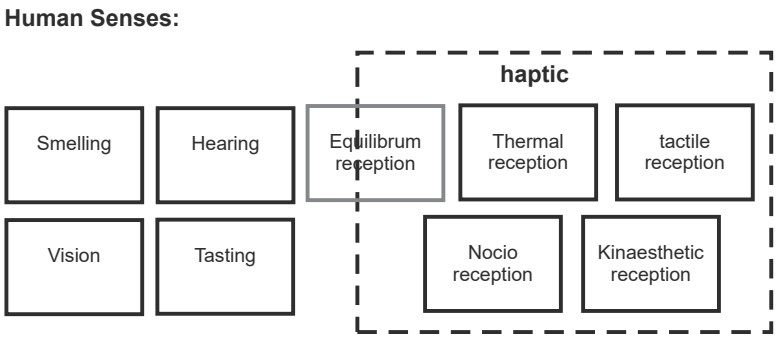


Fig. 2.2 Distribution of senses.

2.2.2 Definition of Haptic Systems

The technical terminology is listed from the special to the general and illustrated by block diagrams. The arrows between the components of the block diagrams may represent different kinds of information depending on the devices they refer to. They remain unlabeled. Haptic devices are capable of transmitting elongations, forces and temperature differences and in a few realizations they also stimulate pain receptors.

> The terms “system” and “device” and “component” are not defined on an interdisciplinary basis. Dependent on one’s point of view the same object can be e.g. “a device” for a hardware-designer ,“a system” for the software-engineer, or “just a component” for another hardware-engineer. These terms are nevertheless part of any engineering discipline and are used accordingly here but should anyhow be read with this knowledge in mind.

A **haptic device** is a system generating an output which can be perceived haptically. It has (fig. 2.3) at least one output, but not necessarily any input. The tactile markers on the keys F and J of a keyboard represent information for the positioning of the index finger. By these properties the keys are already tactile devices. At a closer look the key itself shows a haptically notable point of actuation, the haptic click. This information is transmitted in a kinaesthetic and tactile way by the interaction of the key’s mechanics with the muscles and joints and the force being transmitted through the skin. Such a key is a haptic device without a changing input and two outputs.

A **user** (in the context of haptic systems) is a receiver of haptic information.

A **haptic controller** describes a component of a haptic system for processing haptic information flows and improving transmission. Quite pragmatic in the case of telemanipulation systems these kinds of controllers are frequently either a spring-

damper coupling element between end-effector and the operating element or a local abstraction model of the area of interaction to compensate transmission delays. In the case of a haptic simulator it is quite frequently a simple LTI-model with a high in- and output rate. The LTI model itself is then updated on a lower frequency than the actual speed of the haptic in- and output.

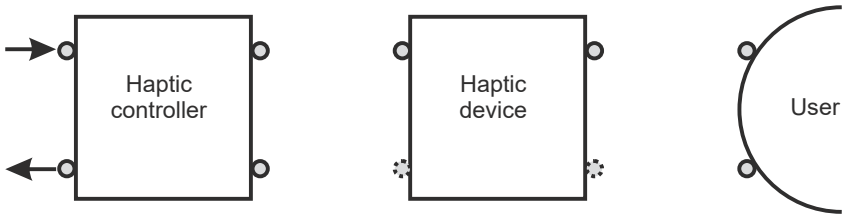


Fig. 2.3 Haptic device, user and controller.

Haptic interaction describes the haptic transmission of information. This transmission can be bi- or unidirectional (fig. 2.4). Moreover, specifically tactile (unidirectional) or kinaesthetic (uni- or bidirectional) interaction may happen. A tactile marker like embossed printing on a bill can communicate tactile information (the bill's value) as a result of haptic interaction.

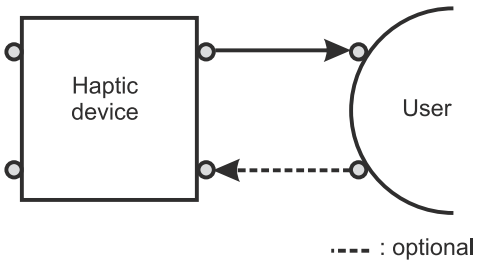


Fig. 2.4 Haptic interaction.

The **addressability** of haptic systems refers to the subdivision (spatial or temporal) of an output signal of a device (frequently a force) or of the user (frequently a position).

The **resolution** of a haptic system refers to the capability to detect a subdivision (spatial or temporal) of an input signal. With reference to a device this is in accor-

dance with the measuring accuracy. With respect to the user this corresponds to his perceptual resolution.

A **haptic marker** refers to a mark communicating information about the object carrying the marker by way of a defined code of some kind. Examples are markers in Braille on bills or road maps. Frequently these markers are just tactile, but there are also kinaesthetically effective ones marking sidewalks and road crossings for visually handicapped people.

A **haptic display** is a haptic device permitting haptic interaction, whereby the transmitted information is subject to change (fig. 2.5). There are purely tactile as well as kinaesthetic displays.

A **tactor** is a haptic purely tactile haptic display generating a dynamic and oscillating output. They usually provide a translatory output (e.g. fig. 9.19), but could also be rotatory (e.g. fig. 2.14).

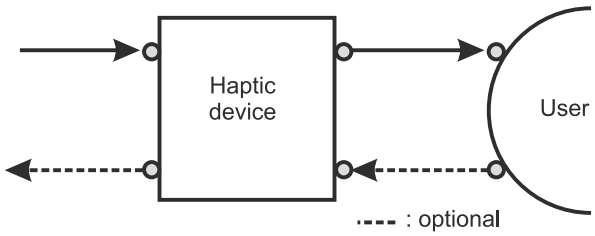


Fig. 2.5 Haptic display.

A **haptic interface** is a haptic device permitting a haptic interaction, whereby the transmitted information is subject to change and a measure of the haptic interaction is acquired (fig. 2.6). A haptic interface always refers to data and device.

Force-Feedback (FFB) refers to the information transmitted by kinaesthetic interaction (fig. 2.6). It is a term coined by numerous commercial products like FFB-joysticks, FFB-steering wheels and FFB-mice. Due to its usage in advertising, the term Force Feedback (FFB) is seldom consistent with the other terminology given here.

A **haptic manipulator** is a system interacting mechanically with objects whereby continuously information about positions in space and forces and torques of the interaction is acquired.

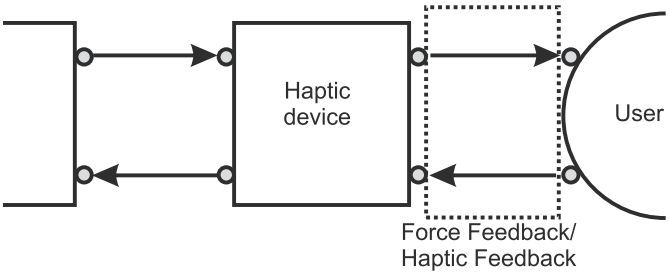


Fig. 2.6 Haptic interface.

A **telemanipulation system** refers to a system enabling a spatially separated haptic interaction with a real physical object. There are purely mechanical telemanipulation systems (fig. 2.7), scaling forces and movements via a lever-cable-system. In the area of haptic interfaces, there are mainly electromechanic telemanipulation systems according to figure 2.8 relevant. These systems allow an independent scaling of forces and positions and an independent closed-loop control of haptic interface and manipulator.



Fig. 2.7 Mechanical telemanipulator for handling dangerous goods (CRL model L) .

A **haptic assistive system** is a system adding haptic information to a natural interaction (fig. 2.9). For this purpose object or interaction properties are measured via a sensor and used to add valuable information in the interaction path. An application would be a vibrating element indicating the leaving of a lane in a drive assistance system.

A **haptic simulator** is a system enabling interaction with a virtual object (fig. 2.10). It always requires a computer for the calculation of the object's physical properties. Haptic simulators and simulations are important incitements for the development of

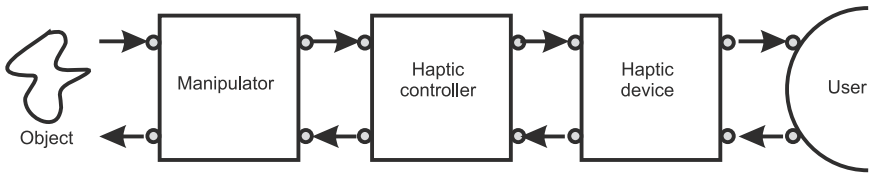


Fig. 2.8 Scheme of an electric telemanipulator.

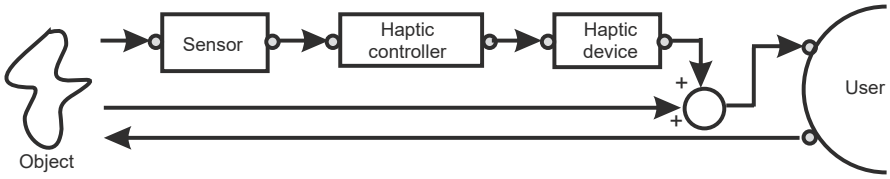


Fig. 2.9 Haptic assistive system.

haptic devices. They can be found in serious training applications, e.g. for surgeons, as well as in gaming applications for private use (see also chapter 13).

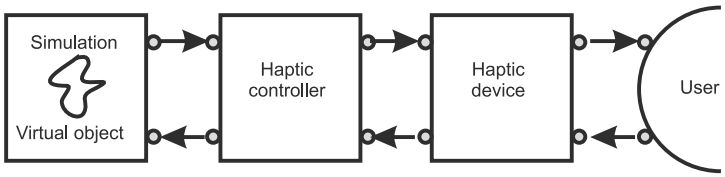


Fig. 2.10 Haptic simulator.

2.2.3 Parameters of Haptic Systems

In [156] LAWRENCE defines the transparency $\underline{T}$ as a factor between impedance as the input source of the haptic interface $\underline{Z}_{\text{in}}$ and the actually felt output impedance $\underline{Z}_{\text{out}}$ of the device.

$$\underline{T} = \frac{\underline{Z}_{\text{in}}}{\underline{Z}_{\text{out}}}. \tag{2.1}$$

The principle of transparency is mainly a tool for control engineering purposes analyzing stability and should be within the range $\pm 3\,\text{dB}$. $\underline{T}$ may be regarded as the sole established, frequency dependent, characteristic value of haptic interfaces. Frequently only the transparency‘s magnitude is considered. A transparency close to “one” shows that the input impedance is not altered by the technical system. The user of the haptic device being the end of the transmission chain experiences the haptic input data in a pristine way. The concept of transparency can be applied to telemanipulation systems and as well as to haptic simulators .

In [39] COLGATE describes the impedance width (Z-width) of a haptic system

$$\underline{Z} - width = \underline{Z}_{\text{max}} - \underline{Z}_{\text{min}} \tag{2.2}$$

as the difference between the maximum load $\underline{Z}_{\text{max}}$ and the perceivable friction and inertia at free space movement $\underline{Z}_{\text{min}}$. The Z-width describes the potential of devices and enables the comparability between them, after technical changes, e.g. by the integration of a closed-loop control and a force measurement.

Active haptic devices are systems requiring an external energy source for the display of haptic information. Usually, these are at least haptic displays. **Passive haptic devices**, on the contrary, are systems transmitting haptic information solely by their shape. This may lead to a false conclusion: A passive system in a control engineering sense is a system with a negative energy flow at its input, e.g. a system not emitting energy into the outside world. This concept of passive control is an important stability criterion which will be discussed in detail in subsection 7.3.3. For the moment, it should be noted that a *passive* haptic system is not necessarily identical with a haptic system designed according to the criterion of *passivity*[1].

The mechanical impedance $\underline{Z}$ is the complex coefficient between force $\underline{F}$ and velocity $\underline{v}$ respectively torque $\underline{M}$ and angular velocity $\underline{\Omega}$. Impedance and its reciprocal value - the mechanical admittance $\underline{Y}$ - are used for the mathematical description of dynamic technical systems. High impedance means that a system is “stiff” or “inert” and “grinds”. Low impedance describes a “light” or “soft” and “sliding” system. The concept of impedance is applied to haptic systems by way of the terms display-impedance or interface-impedance $\underline{Z}_d$. It describes the impedance a system shows when it is moved at its mechanical output (e.g. its handle).The concept of impedance cannot be applied only to technical systems, but also to a simplified model of the user and his mechanical properties. This is described by the term **user-impedance** $\underline{Z}_H$. User-impedance - how stiff a user tends to be - can be influenced at will up to a certain point. Shaking hands can either be hard or soft depending on its frequency. The mechanical resistance of a handshake is lower at low frequencies and

[1] However, in the meaning of this definition, a passive haptic device is indeed a system which is passive, according to the control-engineering classification Yet not all passive systems meeting this criterion of control-engineering are necessarily passive haptic devices.

higher at high frequencies resulting simply from the inertia of the hand's material. Detailed descriptions of the building of models and the application of the concept of user-impedance are given in section 4.2. An introduction into calculating with a complex basis and mechanical systems is given in appendix 16. Understanding complex calculation rules and the mechanical impedances are fundamental to the design of haptic devices in the context of this book. Therefore it is recommended to update one's knowledge by self-studies of the relevant literature of electromechanics [158] and control-engineering [167].

2.2.4 Characterization of Haptic Object Properties

Besides the terminology for haptic systems, there is another group of terms describing solely haptic objects and their properties:

Haptic texture refers to those object properties, which can exclusively be felt by touch. The roughness of a surface, the structure of leather, even the haptic markers already mentioned are haptic textures of the objects they are located on. In some cases a differentiation is made between *tangential* and *normal* textures, whereby the directional information refers to the skin's surface. This specific differentiation is more a result of technical limitations, than of a specialty of tactile perceptions as tactile displays are frequently unable to generate a feedback covering a two or three-dimensional movement.

Haptic shape refers to object properties which can mainly be felt kinaesthetically. This can be the shape of a cup held in one's hand. But it can also be the shape and geometric design of a table rendered to be touched in a virtual environment.

In fact terms like *texture* and *shape* are used analogically to their meaning in graphical programming and software techniques for 3D objects, where meshes provide shape and surface-textures give color and fine structures. However, in comparison with graphical texture, haptic texture mainly describes three-dimensional surface properties incorporating properties like adhesion or friction, i.e. a realistic haptic texture is much more complex in its parameters than a typical graphical texture, even when considering bump-, specular or normal-maps. Therefore numerous **haptic surface properties**, e.g. specific haptic surface effects are defined and described from the perspective of a software engineer. These surface effects are partly derived from physical equivalents of real objects, narrowed down to software-motivated concepts in order to increase the degree of realism of haptic textures:

- **Surface friction** describes the viscose (velocity-proportional) friction of a contact point on a surface.
- **Surface adhesion** Surface adhesion describes a force binding the movement of a contact point to a surface. This concept allows simulating magnetic or sticking effects.

- **Roughness** describes an uniform, sinoid structure of a small, defined amplitude making the movement of a contact point on a surface appears rough.

Tacton refers to a sequence of stimuli adressing the tactile sense. It usually encodes an event within the sequence's pattern. The stimuli vary in intensity and frequency. Both, stimuli and tacton, may even be overlayed with a time-dependent amplitude modulation, such as fade-in or fade-out.

2.2.5 Technical examples

In the following section several technical examples will demonstrate the usage of the terminology introduced before and will give further insight into the interdependency of several terminology groups. The systems presented here should be understood as illustrating examples, not necessarily as technological milestones or devices of extraordinary performance.

2.2.5.1 Force-Feedback Interfaces

There are several commercial haptic control units available on the market for the application in design, CAD and modeling. One major player on the market is SensAble with their PHANTOM®-series and the actually most low-cost product PHANTOM Omni (fig. 2.11a). The PHANTOM-series can most easily be identified by the free positioning of a pen-like handle in a three dimensional space. The position and orientation of this handle is measured in three translational and three rotational degrees of freedom. Depending on the model of the series, the tip force can act on the handle in at least three translational dimensions. The generation of forces is done via electrodynamic actuators; depending on the model these are either mechanically or electronically commutated. The actuators are located within the device's basis and transmit their mechanical energy via levers and Bowden cables on the corresponding joints. As a result of changing level-lengths the transmission-ratio of the PHANTOM devices is nonlinear. For the static situation these changes are compensated within the software driver. The PHANTOM devices are connected to common PCs. The electrical interface used depends largely on the device's product generation and ranges from parallel ports to IDE cards and FireWire connectors. The PHANTOM devices from SensAble are *haptic devices* (fig. 2.11c) primarily addressing the *kinaesthetic perception* of the whole hand and the arm. As the force transmission happens via a hand-held pen tactile requirements are automatically relevant for the design too. This bidirectional *haptic display* is a *haptic interface* to the *user* transmitting force information of a software application in a PC and feeding back positioning information to her or him.

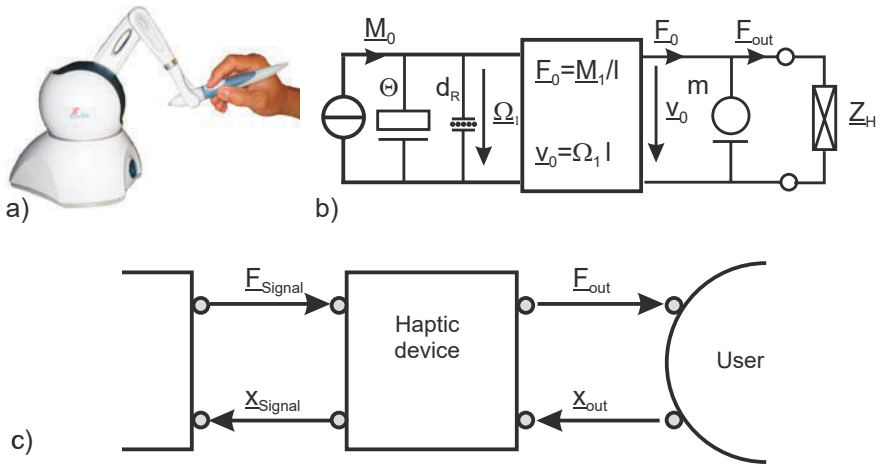


Fig. 2.11 Haptic tool with a SensAble PHANTOM Omni®(a), and corresponding mechanical network of one degree-of-freedom (b), such as block-structure (c).

The network model of one degree of freedom (fig. 2.11b) shows the electronic commutated electrodynamic motor as an idealized torque source $\underline{M}_0$ with inertia of Θ of the rotor and a rotary damping d_R resulting from bearings and links. By the use of a converter resembling levers the rotary movement is transformed in a linear movement with a force $\underline{F}_0$ and a velocity $\underline{v}_0$. An inertia m describes the mass of the hand-held pen. The portion of the generated force $\underline{F}_{out}$ is dependent on the ratio between the sum of all display-impedances against the user impedance $\underline{Z}_H$.

2.2.5.2 Reconfigurable Keyboard

The reconfigurable keyboard (fig. 2.12a) is made of a number of independent actuators arranged in a matrix. The actuators are electrodynamic linear motors with a moving magnet. Each actuator can be controlled individually either as an open-loop controlled force source or as a positioning actuator by a closed-loop control. When being used as force source, the primary purpose of the actuator is to follow a configurable force/displacement curve of a typical key. The application of this reconfigurable keyboard [46] is an alternative to the classical touchscreen - a surface providing different haptically accessible functions depending on a selection within a menu. For this purpose single actuators can be combined to larger keys and may change in size and switching characteristics.

The reconfigurable keyboard is a *haptic device* (fig. 2.12c) mainly addressing the *kinaesthetic* sensation, but has strong *tactile* properties, too. The *user* of the device is the controller of the keyboard, receiving haptic information in form of the changing *shape* of keys and their switching characteristics during interaction. The keyboard is

at least a *haptic display*. As it communicates with another unit about the switching event and the selection, it is also a *haptic interface*.

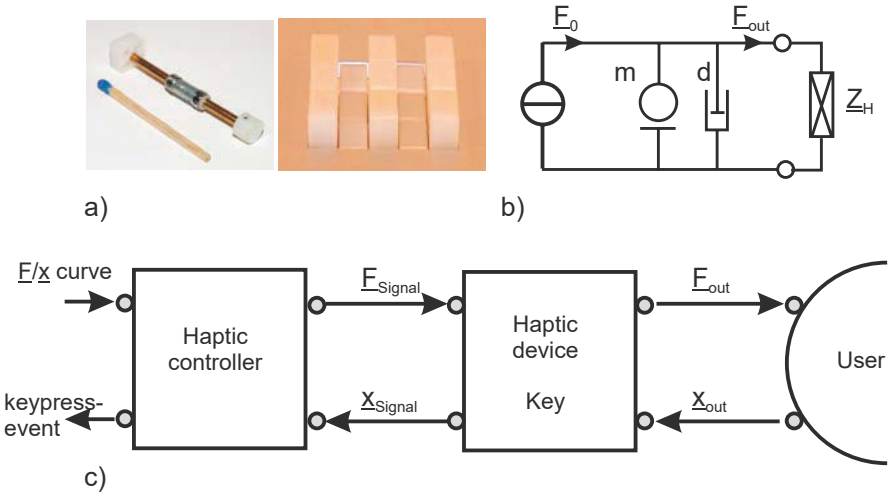


Fig. 2.12 Haptic key as a reconfigurable actuator with a single degree-of-freedom (a), and corresponding mechanical network (b), [46] such as block-structure (c).

The network model (fig. 2.12b) of a single key shows the open-loop controlled force source $\underline{F}_0$ of the electrodynamic actuator, the mass of the moving magnet m and the friction in the linear guides d. Elasticity does not exist, as the design does not contain any spring. This is in contrast to what could be expected from the typical designs of electrodynamic speakers and their membranes. The actuator is capable of generating a force F_{out} dependent on the ratio between the complex impedance of the haptic display $\underline{Z}_D = s\,m + d$ and the user's impedance $\underline{Z}_H$.

2.2.5.3 Tactile Pin-Array

Tactile pin-arrays are the archetype of all systems generating spatially coded information for the haptic sense. Conceptually they are based on Braille-displays whose psychophysical impression has been studied comprehensively since the middle of the 20th century, e.g. by BÉKÉSY [23]. Many approaches were made ranging from electromagnetic actuators of dot matrix printers [232] to piezoelectric bending actuators [149] and pneumatic [292], hydraulic [231], electrostatic [290] and thermal [5] actuators. Tactile pin arrays mainly focus on the skin's stimulation in normal direction. Only lately spatially resolved arrays with lateral force generation are receiving an increased interest [142].

A tactile pin-array with excitation in normal skin direction is a *haptic device* (fig. 2.13c) mainly addressing the *tactile perception*. The *user* is in continuous *haptic interaction* with the device and receives *haptic information* coded in changing pin heights. A tactile pin array is a haptic display. In contrast to the systems examined before this time the user's interaction does not include any user-feedback. As a result the device is *not necessarily* a *haptic interface*[2].

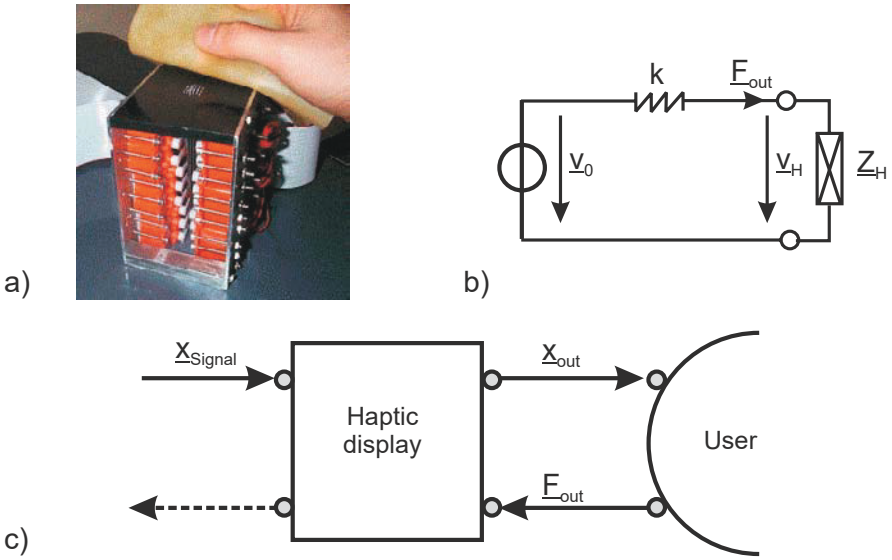


Fig. 2.13 Classical tactile display with pins oriented normal to the display surface [276] (a), and corresponding mechanical network (b), such as its block-structure(c).

In the mechanical network model (fig. 2.13) a tactile pin array corresponds to a positioning or velocity source $\underline{v}$ with a mechanical stiffness k in series to it (a combination of actuator and kinematics). In a stiff design the mechanical admittance of the technical system is small resulting in the elongation being totally unaffected by the user's touch. The system is open-loop position controlled.

2.2.5.4 Vibration-Motor

Vibration-motors are used to direct attention to a certain event. There is a vibration motor similar to figure 2.14a within every modern mobile phone, made of a rotary actuator combined with a mass located eccentrically on its axis. Its rotation speed is controlled by the voltage applied. It typically ranges from 7000 to 12000 rotations

[2] There are possibilities for tactile systems measuring the position of the finger and reporting it to a control unit. But this functionality is not typical of this kind of displays.

per minute (117 to 200 Hz). It is possible to encode information into the felt vibration by varying the control voltage. This is often done with mobile phones in order to make the ring tone haptically perceptible.

A vibration-motor is a *haptic device* (fig. 2.14c) addressing *tactile perception*. The *user* is haptically interacting with the device and receives haptic information in the form of oscillation coded in frequency and amplitude. A vibration-motor is a pure form of a *haptic display*, or more precisely a *purely tactile* display.

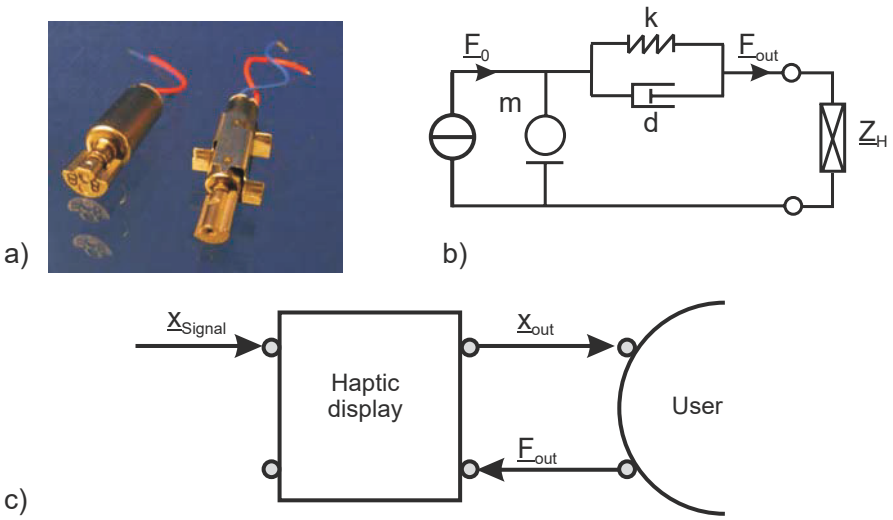


Fig. 2.14 Vibration-motor of a mobile phone (a), and corresponding mechanical network (b), such as block-structure (c).

With vibration motors the relevant force effect is the centripetal force. Assuming a rotational speed of $\omega = 2\pi \frac{10000\,\mathrm{RPM}}{60}$ Hz and a moving mass of 0.5 g on a radius of 2 mm a force amplitude of $F = m\,\omega^2\,r = 1.1\,\mathrm{N}$ is generated, producing a sinoid force with a peak-to-peak amplitude of 2.2 N. This is an extraordinary value of an actuator with a length of only 20 mm. Considering the network model (fig. 2.14b) the vibratory-motor can be regarded as a force-source with sinoid output. It has to accelerate a housing (e.g. the phone) with a mass m which is coupled to the user via an elastic material, e.g. clothes. It is important for the function of the device that the impedance of the branch with spring/damper coupling and user-impedance $\underline{Z}_H$ is large against the mass m. This guarantees that most of the vibration energy is directed to the user, thus generating a perception.

"*This page left intentionally blank.*"

Chapter 3
Biological Basics of Haptic Perception

THORSTEN A. KERN

3.1 The Sense of Touch and its Biology

Examples of haptic systems and the importance of the haptic sense have been discussed in the preceding chapters without actually giving an exact idea of the function of haptic perception. For the design of haptic systems it is vital to have a basic understanding of characteristic biological parameters, as only these will help to identify relevant technical requirements. This chapter introduces the most important terminology and basics for understanding the neurobiology of haptic perception. Please note that research on haptic perception is far from being complete. As a consequence this short presentation of complex biological coherences is a well-founded working hypothesis which will be extended or confuted by further research. In order to perceive information from our surroundings, man is equipped with five senses: Hearing, Smelling, Tasting, Sight and Touch. The physiology of senses distinguishes five sensors and sensory-systems [219] differing from this very popular definition. They allow a classification in a vocabulary lent from a technical approach to describe things:

- **Thermal** sensor for registering the change of temperature especially within the skin,
- **Chemical** sensors reacting on odorous or gustatory substances,
- **Optical** sensors reacting on the impact of photons, especially within the cones and rods in the retina,
- **Pain** sensors, also named nociceptors, to identify chemical and physical tissue damage,
- **Mechanical** sensors for detecting mechanical tensions and strains e.g. within the skin or muscles.

T.A. Kern (ed.), *Engineering Haptic Devices*, DOI 10.1007/978-3-540-88248-0_3,
© Springer-Verlag Berlin Heidelberg 2009

The sensory capacity and its importance for haptic perception are valued differently. The visual sensors register ≈ 10 Mio. bit/s, the sense of touch ≈ 1 Mio. bit/s and the acoustic sense ≈ 100 kbit/s [18]. The processing of these sensory data happens within the cerebral cortex. It is structured in functional brain areas. The primary motor cortex is the physiological location for processing data from the sense of touch. A visualization of the distribution of body parts on the primary motor cortex (fig. 3.1) shows a significant portion being used for fingers and hand.

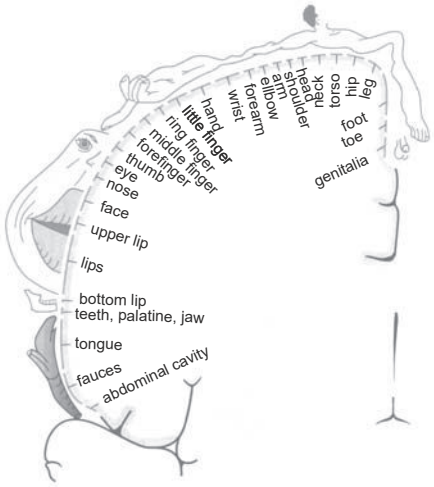


Fig. 3.1 Visualization of the functional brain areas in the motor cortex (somatotrope ordering) [52].

Within the sensorimotor functions the haptic sense has the highest importance. It consists of a group of mechanical sensors detecting force induced deformations within tissues in the skin, muscles and joints. As a consequence haptic perception is the sum of signals from a large number of measurement points distributed among the human body, consisting of at least 6 types of sensors which can be divided into two basic groups: Tactile and kinaesthetic sensors (fig. 3.2).

Tactile sensors are located in the outer areas of the skin in exposed positions (e.g. the fingertips). They react on strains of the skin and are activated either proportionally to the elongation, to the velocity or to the acceleration. The neuro-pyhsiology distinguishes between four different types of tactile sensors [236, 219]:

- Rapid-adaption or Fast-Adaption (RA or FA-I) *Meissner corpuscles* - with velocity dependent activation.

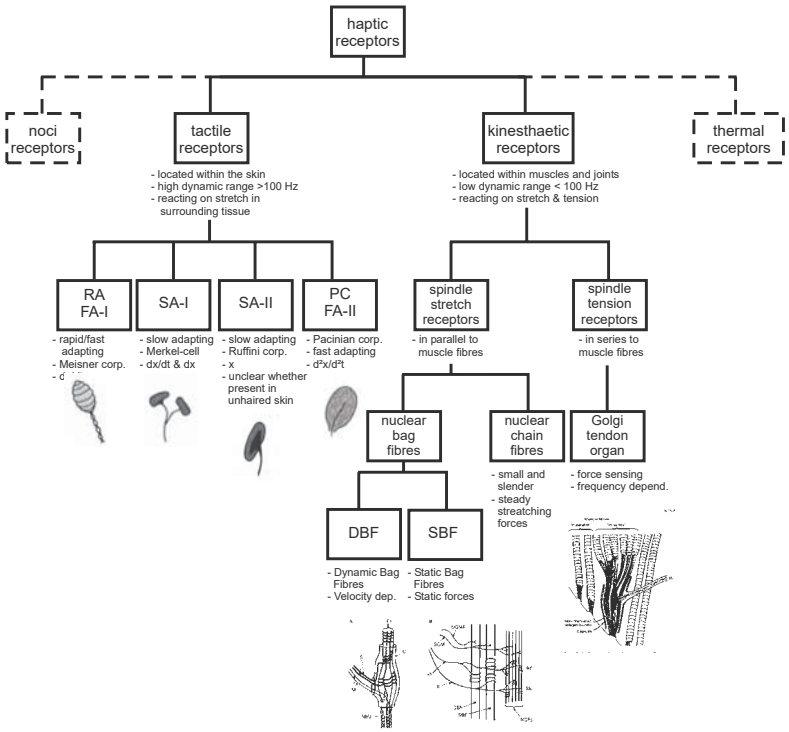


Fig. 3.2 Classification of the haptic receptors and their naming, adapted from [236].

- Slow-adapting (SA-I and SA-II) *Merkel cells and Ruffini-corpuscles* with velocity-dependent and elongation proportional activation. They show a lower dynamic response [1] compared to the Meissner-corpuscles.
- Fast-Adaption (FA-II) *Pacinian corpuscles* with acceleration-proportional activation.

The distribution of sensors varies within different skin areas (fig. 3.3) and is part of current research. For example in [193], the existence of Meissner corpuscles has to be put into question in contrast to established doctrines.

Unlike tactile sensors kinaesthetic sensors are located mainly within muscles, joints and tendons. They acquire forces affecting whole extremities only. Their dynamic requirements are reduced as a result of the mechanical low-pass characteristics of the extremities (their mass, damping and stiffness). Their requirements on the relative resolution between the smallest perceivable force vs. the maximum de-

[1] *Dynamic* refers to the width in frequency domain the actuator or sensor is designed for. There is an important difference to amplitude-dynamic referring to the height of the output from an actuator or the maximum input value of a sensor.

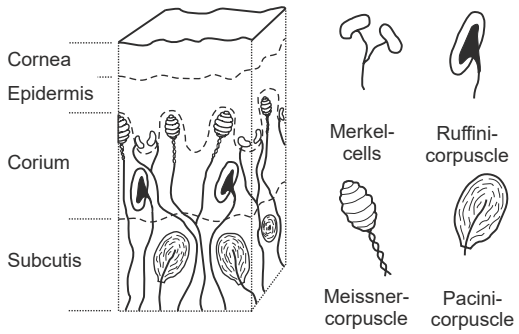


Fig. 3.3 Distribution of the sensors in the skin layers [46].

tectable force (amplitude-dynamics) can be compared with the group of tactile sensors. Kinaesthetic sensors can be divided in two groups:

- spindle-stretch-receptors *Dynamic Bag fibres and Static Bag fibres* placed in parallel to the muscle fibres.
- spindel-tension-receptors *Golgi tendon organ* - in serial orientation to the muscle fibres.

When summarizing all information about biological sensors contributing to haptic perception, it is interesting to see that nature chooses a design to identify forces and vibrations, which does not significantly differ from the technical solution of comparable problems. However comparable technical solutions are older than the biological understanding of the sense of touch. Therefore it seems likely that with the given physical constraints only solutions optimized in such a manner are adequate.

Besides dealing with theses sensors as the first part of the haptic perception chain, in the next step it is necessary to consider a model for the neurological processing of haptic information, in order to get a feeling for the complexity of the system and outline components relevant for the design of technical haptic systems.

Figure 3.4 shows a simplified understanding of neuronal subsystems participating in a task like “grasping a glass of water”. The *motivation phase* starts with thirst due to e.g. the body salinity being too high, and with the knowledge about the availability of a glass of water. As a result a decision is taken to “seize the glass of water”. Within a *programming phase* this decision results in a definition of movements for single extremities and body parts. As a subcomponent each of them has a position controller, which controls the movement. Feedback is given by the motor sensors within the joints but also by visual control from a superordinate control loop. Subordinate to the visual control, a closed-loop circuit with force feedback exists, enabling the safe and secure holding of the glass based on a maximum force to be exerted. As an alternative, a feedback loop can be assumed controlling the grasping force to avoid a slipping through of the glass.

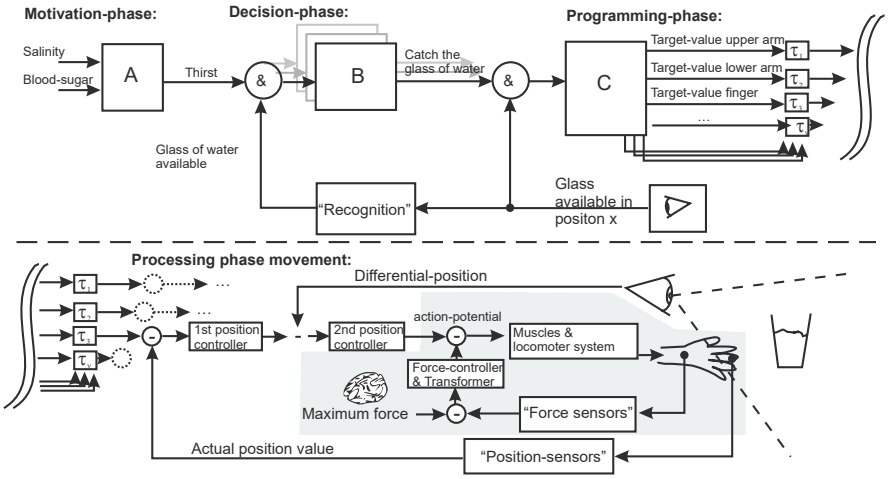


Fig. 3.4 Neuronal processing chain with the example of grasping a glass of water:
A, B, C are decision elements of no further detailation; τ_i delay element.

It is remarkable that the analogy to technical control systems can so easily be drawn. Decision phase, programming phase and processing phase are accepted references for central-neural structures [219]. The interconnection between position and force-controller is a direct result from dynamic ranges and measurement errors unique to the components of the close-loop circuit. The position control loop including the locomotor system and kinaesthetic sensors shows a dynamic range of $\leq 10\,\text{Hz}$ [287, 83]. Additionally angle-positioning and absolute position measurement without a line of sight show large errors (2° to 10° dependent on the joints participating [34]). Movements including visual control are much more precise. The visual perception is able to resolute movements with up to 30 Hz depending on the illumination level. By the aid of sight a human is able to move to a position and hold it until immediately before physical contact - which is strictly impossible with closed eyes. On the other hand tactile sensors show a dynamic range of many hundred Hertz. This capability combined with high amplitude-dynamics enables humans to hold even slippery and fragile objects without breaking them.

3.2 Haptic Perception

Knowledge about the performance of haptic perception is essential. for the formulation of requirements as a basis for system design, For each body part there are different characteristic values, as the haptic sense is not located in a single organ. Additionally haptic interaction is always bidirectional, which means - especially in case of kinaesthetic interaction - that haptic interaction can be mapped only by

considering positions and angles of body parts and forces and torques of mechanical specifications,. Furthermore haptic perception is also greatly dependent on the dynamic excitation in a broad frequency range. Last but not least the aspect of multimodality[2] has to be considered. Haptically inconspicuous keys and buttons are considered to be of high quality when they are accompanied by a loud click-sound, compared to silent buttons with identical haptic properties. As a consequence of the complexity of effects influencing haptic perception every characteristic value taken from literature has to be seen in the context of the individual test design and weighted with the accuracy of the experiment's layout. The characteristic values presented within this chapter shall be taken as points for orientation only, and should be modified and even disapproved of by future experiments.

3.2.1 Psychophysical Concepts

In order to be able to understand characteristic values of haptic perception, a basic knowledge about some relevant psychophysical concepts is necessary. The definitions given here are based on G. A. GESCHEIDER [66], recommended to any reader interested in this subject.

3.2.1.1 Threshold and Difference-Limen

Two fundamental concepts for the analysis of thresholds are distinguished in psychophysics. On the one hand there is the measurement of thresholds for differential perception (thresholds of differential sensitivity). On the other hand there are thresholds for absolute perception (thresholds of absolute sensitivity). All measurement principles in psychophysics can be categorized according to these two principles. Additionally the analyzed stimuli differ in their dimensions (e.g. space, time, spectral [3]).

The **absolute threshold** (fig. 3.5) of a stimulus describes the value, from which a stimulus ϕ begins to become perceivable.

As another characteristic value the stimulus' change is relevant, creating a **just-noticeable difference (JND)** [4]. The stimulus' change is called **difference threshold** or alternatively **difference-limen (DL)**. Consequently the DL means the measurement of a $\Delta\phi$ being the difference to a stimulus ϕ_0 compared to another stimulus ϕ_1. The JNDs are numbered discretely as JND being a member of N. The first JND is the first DL after the absolute threshold; the second JND is that DL following the sum of the absolute threshold and the first DL (fig. 3.6). To sum up: the JND is the

[2] Multimodality: Cumulated influence of different sensory perceptions affecting the perception of a single event.

[3] of different frequency

[4] rarely also named **just-noteable-difference**

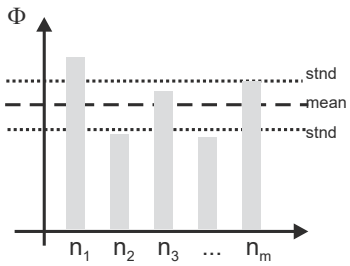


Fig. 3.5 Identification of the absolute threshold of a stimulus ϕ from m tests .

smallest physiological scale unit of the linearized perception of a physical stimulus ϕ.

$$\mathrm{DL} = \Delta\phi = \phi_0 - \phi_1 \tag{3.1}$$

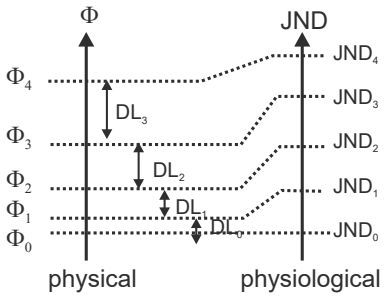


Fig. 3.6 Concept of the physical scale of the stimulus ϕ and the linearized physiological scale of the discrete JND.

The DL of stimuli is analyzed according to different questions. The methods applied allow conclusions concerning the neuronal processing of stimuli. A classical method for the analysis of DL is the presentation of a reference stimulus comparing it to a second stimulus, which is presented to a subject either in an automated way or manually controlled by the test-person himself (fig. 3.7).

Besides the aspects just mentioned there are others for doing comparable analysis. On the one hand, there is the aspect of **masking** with the question: “At which point will two stimuli dependent on a single parameter be perceived as different?”. Aspects analyzed frequently for masking a time and spatial dependencies. As a result the terms of **temporal masking** and **spatial masking** have been fixed.

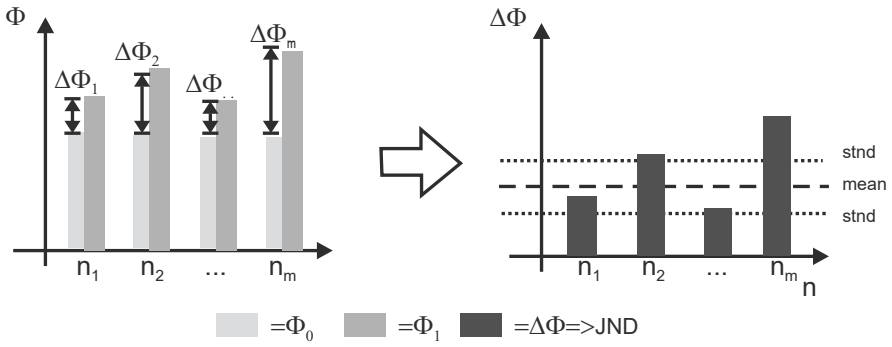


Fig. 3.7 Identification of the DL as a result from the Just-Noticeable-Difference (JND) $\Delta\phi$ between the stimuli phi_0 and phi_1 from m experiments.

example dynamic masking: The perception of a change in frequency of a mechanical oscillation of fixed amplitude shall be analyzed. For this purpose two stimuli are given at the same time to test a subject. The subject is allowed to change the frequency of one stimulus until he or she detects two independent stimuli. The measure of the change in frequency Δf is the value for DL with respect to the reference stimulus. Results of such a kind of experiments are not always precise and should be critically analyzed. For example in case of stimulus locations very near to each other the above experiment can be easily interpreted in a way that only the maximum amplitude of a summed up signal has been analyzed and not the DL of a frequency change. To prevent this kind of criticism a careful experiment design should be done with a series of hypotheses for falsi- and verification. In this case an additional experiment would be adequate showing a statistically significant difference in the perception of JND between a summed-up amplitude of two stimuli with identical frequency compared to a signal with two frequencies.

example temporal masking: A stimulus ϕ_0 with a frequency f is presented for a long period t. Afterwards stimuli ϕ_n, e.g. dependent on frequency, are given. The perception of those stimuli (e.g. with regard to the absolute threshold) varies dependent on the prior period t. The measure of this variation is the temporal-masking-effect of a certain masking frequency f.

example spatial masking: Two stimuli ϕ_0 and ϕ_1, e.g. needles on the skin are given with a spatial distance d. At a certain distance d both stimuli are perceived independently from each other. This is a very specific example of spatial masking frequently used for measuring the resolution of tactile perception. It has therefore been given its own term: **two-point-threshold**

Another aspect of analysis is the **Successiveness Limen (LM)** connected with the question: “How many stimuli presented consecutively can be perceived?”

> **Example LM:** With the help of a vibratory motor a sequence of stimuli is presented on a body location. The stimuli vary according to a temporal pattern. The LM is the temporal pace enabling a correct perception of the sequence.

3.2.1.2 Psychophysical Laws

An important way for presenting DL $\Delta\phi$ is as a value related to a reference stimulus ϕ_0 according to the formula

$$\frac{\Delta\phi}{\phi_0} = c \tag{3.2}$$

In 1834, E.H. WEBER found out, that c is a constant quotient for a specific perception. In his key experiment he placed weights on the skin and found out, that c is almost $\frac{1}{30}$. This means that the next higher weight differing from a weight of 200 g is $\frac{1}{30} \cdot 200\,\text{g} + 200g = 206.66\,\text{g}$. The value c differs significantly between different stimuli, but the comprehensive coherence according to equation 3.2 (Weber’s law) seems to be universal for many situations. As a consequence Weber’s law allows putting different senses and their perception in relation to each other. An exception is the area of lower stimuli (fig. 3.8a) in the range of absolute thresholds where c increases significantly.

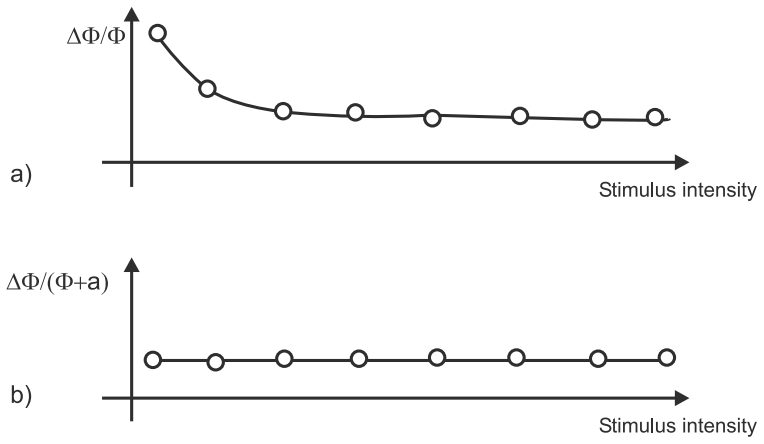


Fig. 3.8 Identification of the Just-Noticeable-Differenz (JND) $\Delta\phi$ in between the stimuli phi_0 and phi_1 of m experiments.

A modification of *Weber's law*

$$\frac{\Delta\phi}{\phi_0 + a} = c \tag{3.3}$$

compensates this dependency in the range of absolute thresholds (fig. 3.8b). The constant a is - identical to c - specific for each sense and as compared to c, quite low. The physiological reason for a has not finally been determined. A hypothesis existing assumes it to be a measure for the background noise of the corresponding receptors.

Some senses, especially the acoustic but also the haptic sense, show a nonlinear logarithmic dependency on perceived intensity and physical excitation. For the range of stimuli, for which *Weber's law* is valid according to its original formula (equ. 3.2) a new dependency can be formulated. This dependency, named **Fechner's law**,

$$\Psi = k \log \phi \tag{3.4}$$

provides a linearized measure Ψ of the perception amplitude.

Today *Fechner's law* has mainly a historical significance. In 1975 it was replaced by S.S. STEVENS, suggesting a law describing the intensity of a stimulus by an exponential relation:

$$\Psi = k\phi^a \tag{3.5}$$

This relation is called **Power-law** and allows comparisons of numerous perception-dependencies by a look at its constants a and k. If $a = 1$, the equation 3.5 gives a linear dependency. At values for $a > 1$ the law gives a dependency increasing with increased stimulus, at $a < 1$ a damping of the perception with increased stimulus is resulting. When logarithmizing equation 3.5, an interdependency easy to display on diagrams with logarithmic axis (fig. 3.9) can be obtained.

$$\log \Psi = \log k + a \log \phi \tag{3.6}$$

with y-axis $\log k$ and a slope of a.

Table 3.1 gives an extraction of STEVENS' published data [242] of the coefficient a according to equation 3.5.

3.2.1.3 Mean Values and Percentiles

The analysis of psychophysical measures is always laborious due to large variances in results either between individual subjects or as in certain tests, among specifically trained test-persons. As a consequence, statistical design and the application of signal detecting algorithms should be considered for any such experiments. For details

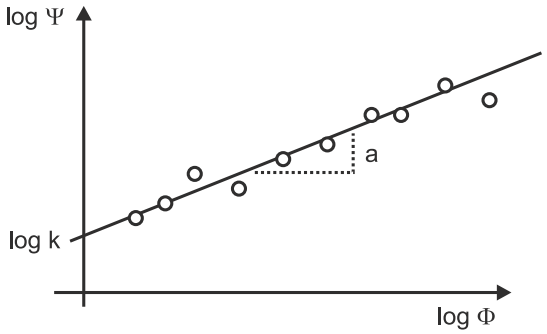


Fig. 3.9 Logarithmic display of hypothetic measurements according to equation 3.6 [66].

Table 3.1 Exponents *a* of the power function 3.5 relating sensory magnitude to stimulus intensity - extract taken from [242]

Continuum	Measured exponent	Valid for stimulus
Sound level	0.67	Acoustic pressure of a 3 kHz signal
Vibration	0.95	Oscillation with 60 Hz of unknown ampltiude at the index finger
Vibration	0.6	Oscillation with 250 Hz of unknown ampltiude at the index finger
Taste	1.3	Sugar
Taste	0.8	Sweetener
Coldness	1.0	Metal contact at the arm
Warmness	1.6	Metal contact at the arm
Muscular strength	1.7	Static contraction
Pressure on the palm	1.1	Static force on skin
Finger span	1.3	Thickness of blocks
Angular acceleration	1.4	5-sec rotation
Electrical shock	3.5	Current through finger
...	...	...

of these procedures, literature of statistical experiment design and [66] is suggested. For a more general perspective the following remarks should be considered:

Frequently, in psychophysics experimental results follow a Gaussian normal distribution. This happens with regard to a single person as well as with respect to a larger number of people. Of course, a **Gaussian distribution** can be characterized by a mean-value μ and a standard deviation σ. The mean value defines the measure where exactly 50% of a given set (e.g. of experiments) are above and below that value.

For the usage with sets not following a normal distribution, the usage of percentiles is suggested. Typical examples of their application are anthropometric values in ergonomics. The x-th **percentile** gives the point on a scale, where *x* percent

of tests of a given set are below that value. In the exceptional case of a normal distribution the fiftieth percentile is identical to the mean-value (fig. 3.10).

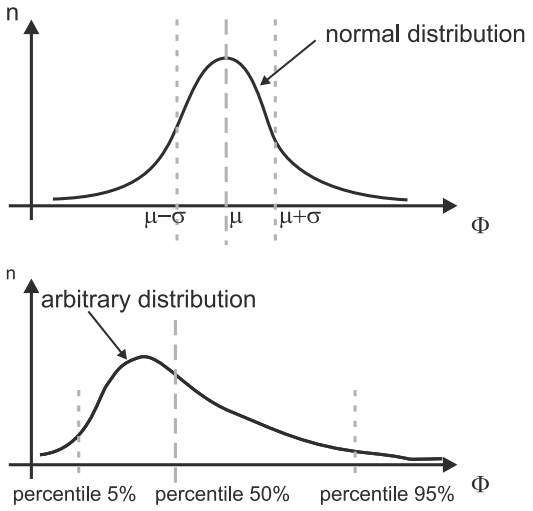


Fig. 3.10 Gaussian distribution with medium value μ and standard deviation σ compared with an arbitrary distribution characterized by percentiles.

3.2.2 Frequency Dependency

As mentioned in section 2.2, every kinaesthetic interaction has a tactile component. We know from the analysis of "grasping a glass of water" in section 3.1 that tactile components are part of the interaction's innermost feedback-loop. As a result the requirements for their dynamic properties are extraordinarily high. This section discusses the perception thresholds and difference-limens as identified in neurology and psychophysics from an engineering perspective. It is therefore a preparation for the identification of requirements for technical systems interfacing the sense of touch.

The identification of haptic perceptional dynamics can be performed either with psychophysical methods or with neurological tools. When focussing on the receptors only, the analysis of tactile and kinaesthetic sensors can be done independently from each other. Neuronal potential on nerve fibres can be measured via interventional implanted electrodes, and even positioned during electrode recording [250, 131].

In [122] several tactile sensory types (compare fig. 3.2 on page 37) have been analyzed as to their frequency-dependency (dynamics) and their thresholds for the detection of skin deformation (fig. 3.11). Frequency areas of slow-adapting (SA) and rapid-adapting (RA) sensors complement and overlap each other. The SA-II sensor especially affects a range of $\approx 8\,\mathrm{Hz}$. According to this study the mean threshold of

the isolated sensors shows a maximum in sensitivity at around $\approx 300\,\mathrm{Hz}$ with an elongation of $10\,\mu\mathrm{m}$.

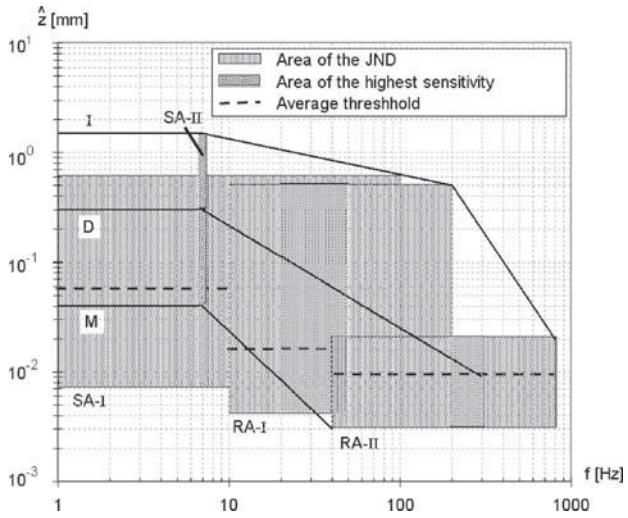


Fig. 3.11 Frequency dynamics and thresholds for the detection of skin deformations according to [122], summarized by [119] .

WILKINSON performed a study [287] on thresholds on isolated Golgi-tendons receptors (fig. 3.12). The results show an almost linear dependency between the response of receptors in mV and the stimulation in μm over frequency. The relevant frequency range of these receptors is lower than the range of the tactile receptors of figure 3.11, especially as the masses and stiffnesses of limbs show distinct low-pass characteristics. High frequency components of forces and elongations are damped anyway and therefore the kinaesthetic sensory system does not have to be able to measure it.

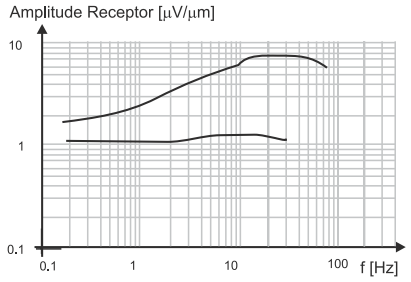


Fig. 3.12 Frequency dependent thresholds of isolated Golgi-tendon-receptors [287].

For the design process of haptic systems the focus point does not lie on single biological receptors but rather on the human's combined perception resulting from the sum of all tactile and kinaesthetic sensors. In this area numerous studies have been performed, three of which are given here showing the range of results. In 1935, HUGONY already published a study about the perception of oscillations dependent on frequency of mechanical stimuli [101]. Additionally he quantified different stimuli-levels which are defined from the absolute perception threshold to pain-thresholds (fig. 3.13a). To complement this general study TALBOT added details about the interdependency of the isolated biological sensor and perception (fig. 3.13b). Both scientists showed that the sensitivity of perception increases to a frequency of $\approx 200\,\mathrm{Hz}$ (HUGONY) and $\approx 300\,\mathrm{Hz}$ (TALBOT). These two studies along with several others were compiled by HANDWERKER [219] resulting in a combined curve of haptic perception thresholds (fig. 3.14).

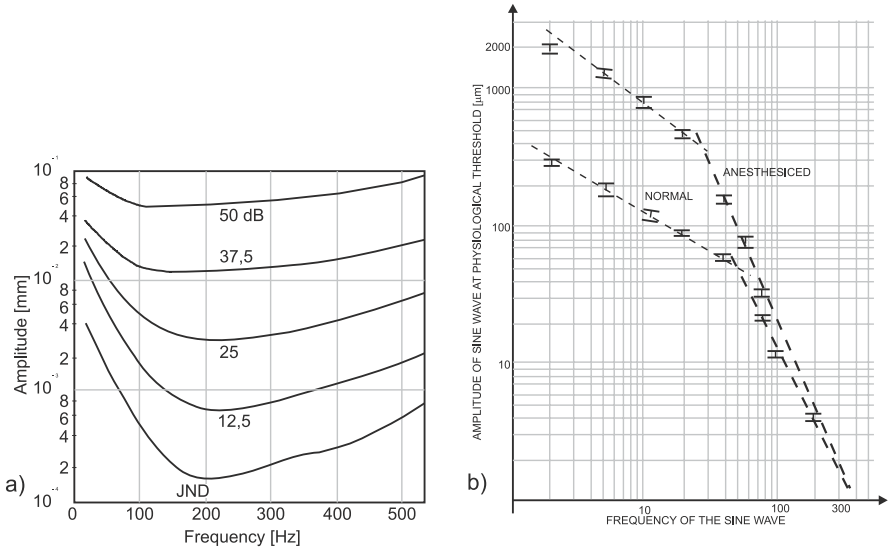


Fig. 3.13 Thresholds for the perception of oscillations after HUGONY 1935 [101] (a) and study from TALBOT about the perception of oscillations [249](b).

A source for the analysis of haptic perception worth to be recommended can be found in the publications by GESCHEIDER. He followed a stringent analysis-methodology and discussion of haptic sensory systems. Beginning in 1970 until at last in 2002 a series of hypotheses and measurements in numerous publications is documented. Another source worth to be considered is the work by BÉKÉSY [23] and by JOHANSSON.

Next to the already known dependency of haptic perception on frequency, another dependency exists connected to the surface area transmitting the mechanical oscillations: Large areas of force transmission ($A > 1\,\mathrm{cm}^2$) and small areas of force transmission ($A < 1\,\mathrm{mm}^2$) differ significantly according to their absolute perception

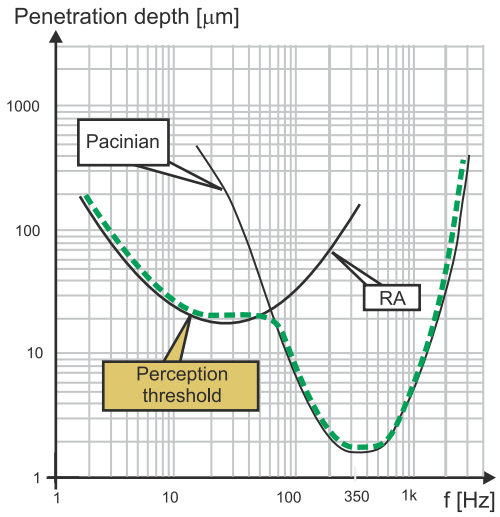


Fig. 3.14 Combined curve of thresholds of haptic perception [219].

thresholds (fig. 3.15). When focusing on kinaesthetic devices, usually a large area of force transmission exists. With tactile devices smaller force transmission areas have to be considered. The perception curve is a combination of the four tactile sensor types and shows a minimum (point of maximum sensitivity) at $\approx 350\,\text{Hz}$. The frequency-dependency is obvious and undoubted. Only the precise shape and the exact position of the minimum vary in the range of $\approx 100\,\text{Hz}$ depending on measurement, author and publication. Additionally it can be noted that the perception of very low frequencies below 0.1 Hz was not subject to many studies. Typically the perception curves are assumed to stay constant to lower values from a frequency of approximately 1 Hz.

Besides frequency-dependency there exist two additional dependencies affecting haptic perception. Ongoing mechanical stimuli result in a reversible desensitization of receptors. This time dependency is used in [69] to mask single receptor classes in order to study the part of other receptor classes in overlapping frequency areas. The time dependency of perception curves ΔK in dB can be approximated according to

$$\Delta K(t) = 12 \cdot (e^t)^{12}. \tag{3.7}$$

As a result desensitization happens in a time frame of a second (spectral components below 10 Hz). As a consequence desensitization is not a matter to be necessarily considered for the design of haptic devices, telemanipulation systems or simulators due to their large ratio between usage vs. desensitization time frame. A steady state can be considered for almost all relevant applications. In practical application this approximation is not necessarily adequate. For example when tactile devices based on pin- or shear-force systems are used, there is some evidence, according to the author's purely subjective observation, that the mentioned effect still happens after minutes of usage.

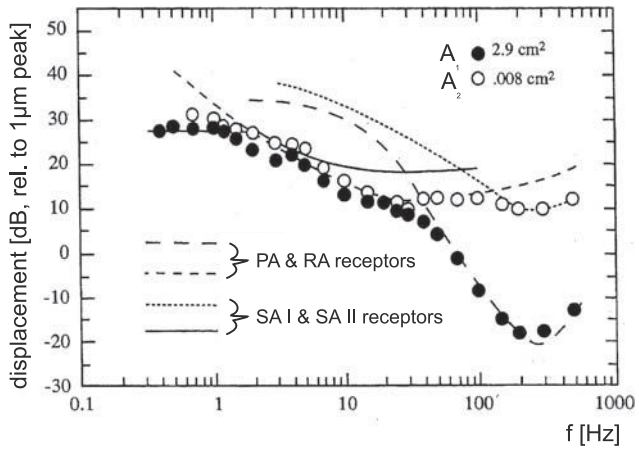


Fig. 3.15 Frequency-dependency of the perception of oscillations, acquired with two different sizes A_1 and A_2 of the contact area [67] (note: *PA-, RA-, SA-receptors from fig. 3.2*).

The amplitude-resolution (DL) of haptic perception shows a logarithmic dependency analogue to the visual and acoustic perception. The perception of smallest changes dependent on frequency with varying base excitation was studied in [68]. Measurements were taken at two frequencies (25 Hz, 250 Hz) and with white noise. The approved dependency of DL of the amplitude of the base excitation is nonlinear with a maximum difference of $\approx +3$ dB. It is larger for smaller amplitudes of the base excitation. This allows the conclusion that the *Power-Law* (section 3.2.1.2, equ. 3.6) can be used for the description of perception. However its coefficients have to be identified for every contact situation independently.

3.2.3 Characteristics of Haptic Interaction

Besides the dynamics' curves in the prior section there are numerous insular values from experiments documenting the possibilities of haptic interaction. The results can be divided into two groups. In the table of *haptic perception* (tab. 3.2) the parameters from a receptive perspective are summarized. In the table of *active movements* (tab. 3.3) border values of the capabilities of the active parts of motor systems are summarized. The tables are based on a collection by DOERRER [46] and have been extended by selected additional sources. However, when considering their application, a very important statement of BURDEA [34] still has to be remembered: "... that it is dangerous to bank on recommendations for the design of haptic devices, especially when they are taken from different experiments, with varying methods, and when only a small number of participants took part". The characteristic values

given here can only represent a selection of the analyses presented in literature. For quite an actual and a very compelling summary [118] is recommended.

3.2.3.1 Characteristic Values of Perception

Table 3.2 Characteristic values of the perceptive capabilities of the human hand.

Base item	Characteristic value	Body part	Value	Source
Static elongation / position	Skin-deformation, absolute value [(a)]	Fingertip (tactile)	10 μm [(b)]	[123]
	Two-point threshold [(c)] (Spatial resolution)	Fingertip (tactile)	2-3 mm [(d, e, f)]	[34][123]
		Palm (tactile)	10-11 mm	[123] [230]
	Position-resolution, Difference-limen (DL)[(g)]	finger joint (kinaesthetic)	2.5 °	[252]
		Wrist (kinaesthetic)	2.0 °	[252]
Dynamic elongations (vibration)	Frequency, upper limit (tactile perception)	Finger (tactile)	5-10 kHz	[31][34]
	frequency, upper limit (kinaesthetic perc.)	whole body (kinaesthetic)	20-30 Hz	[31]
	Maximum sensitivity	Fingertip, palm (tactile)	at 200-300 Hz	[15][25] [123]
	amplitude, absolute threshold	Fingertip, palm (tactile)	0.1-0.2 μm at 200-300 Hz [(h, i, j)]	[15][25] [123]
	Amplitude-resolution, difference-limen (DL)[(g)]	fingertip (tactile)	10-25 %	[25]
	Frequency-resolution, difference-limen (DL)	Fingertip (tactile)	8-10 % [(k)]	[25]
Force and pressure	Force, absolute threshold	Fingertip [(l)] (tactile)	0.8 mN	[34]
		Palm (tactile)	1.5 mN	[34]
	Force, Difference-Limen (DL)	Total body (kinaesthetic)	5-10 % (ca. 7 %) [(m, n, o)]	[192]
	Pressure, Absolute threshold	Finger (tactile)	0.2 N/cm^2 [(p)]	[229]
	Pressure, Difference-limen (DL)	Wrist (kinaesthetic)	4-19 % [(q)]	[252]
Torque	Difference-limen (DL)	Thumb, index finger (kinaesthetic)	12.7 % [(r)]	[114]
Elasticity	Difference-limen (DL)	Thumb, index finger (kinaesthetic)	5-15 % [(s, t)]	[251]

Annotations and additions to table 3.2

(a) absolute threshold: starting from this threshold a stimulus is perceived.
(b) If movement is permitted, surface structures of 0.85 μm height can be perceived [141].
(c) Two stimuli with a spatial distance larger than the two-point threshold can be distinguished from each other. The spatial resolution is the reciprocal value of the two-point threshold. The table of two-point thresholds at the human body are compelling. In a general the values increase from the fingertip at 1 mm up to belly and back with 10 mm. The importance of two-point threshold is discussed contradictory, as the perception of shear forces is getting more attention today. Nevertheless the two-point threshold gives a good qualitative impression of the receptor density on the body surface.
(d) With vibrotactile stimulation the two-point threshold is at the lower limit of 2 mm [123].
(e) The two-point threshold decreases, if the two stimuli are presented short after another [123].
(f) A position change of a stimulus can be resolved spatially up to ten times better than the static two-point threshold [123].
(g) see section 3.2.1.1
(h) The perception threshold is strongly dependant of the vibration frequency, the location of the stimulus and the size of the contact area [25][123][230].
(i) Amplitudes larger than 0.1 mm are perceived as annoying at the fingertips[25].
(j) A stimulation with constant frequency and amplitude results in a desensitization, increasing up to a numb feeling which may last several minutes after the end of the stimulation. [35][122].
(k) The capability to differ stimuli is reduced after 320 Hz [31].
(l) The absolute-threshold of the force perception is perceived by tactile mechanoreceptors transmitted by deformations of the skin.
(m) The value of 7 % was identified almost independent of the muscle-system and the absolte value of the force [192].
(n) A special experiment with forces between thumb and index finger gabe a JND of 5-10 % for reference forces between 2.5 and 10 N at a finger-distance of 45 to 125 mm. For smaller forces a larger JND is expected [192].
(o) In an experiment to distinguish masses of different objects JNDs of 10 % were identified [15].
(p) The tactile perception of the human is very sensitive for pressure gradients and especially for edges [252].
(q) Experiment made at a reference pressure of 1.8 N/cm^2. JND increased strongly with reduced contact area: 4.4 % at 5.06 cm^2, 18.8 % at 1.27 cm^2 [252].
(r) Experiment with a reference torque of 60 mNm.
(s) Experiment with a reference elasticity of 4 mm/N and a spring elongation of 15 to 35 mm.
(t) The mechanical work lost during compression and deformation of object surfaces is relevant for perceiving elasticity [34].

3.2.3.2 Characteristic Values of Movement

Table 3.3 Characteristic values of the output capabilities of the human hand.

Base item	Characteristic value	Body part	Value	Source
Velocity	Maximum velocity	Finger (joint nearest to the palm)	17 rad/s [(a)]	[87]
		Finger (middle joint)	18 rad/s [(a)]	[87]
Movement	Bandwidth (Movement and force)	Finger	5-10 Hz [(b)]	[31]
		Arm, unexpected reaction	1-2 Hz	[239]
		Arm, periodic movement	2-5 Hz	[239]
		Learned trajectory	bis 5 Hz	[239]
		Reflex	bis 10 Hz	[239]
Force	Maximum force	Index-, middle-, ringfinger	40-50 N [(c)]	[34]
		Thumb	85-100 N	[87]
		Wrist	35-65 N	[34]
		between 2 fingers: fingertip-grasp [(d)]	45-65 N [(e)]	[34]
		between 2 fingers: key-grasp [(f)]	76-109 N	[34]
	Movement precision	Fingertips (with visual feedback)[(g)]	40 mN [(h)]	[239]
		Fingertips (without visual feedback)	11-15 % [(i)]	[239]
Torque	Maximum torque	Finger (palm-nearest joint)	370-500 Ncm [(j)]	[87]
		Finger (middle joint)	120-289 Ncm [(j)]	[87]
		Finger (most distal joint)	40-85 Ncm [(j)]	[87]

Comments and annotations to table 3.3

(a) For slow movements velocities of 3-6 rad/s for palm-nearest and middle finger-joints are documented [87].
(b) The bandwidth of the ouput capacity is situation dependent: unexpected signas: 1-2 Hz, periodic signals: 2-5 Hz, known movement patterns: ca. 5 Hz, reflex-reactions: approx. 10 Hz [31].
(c) With stretched fingers and movements only at the palm-nearest joint. These values are quite independent from the actual joint position (flection in between 0 and 80 °) [34].
(d) Several grasp-positions of the finger have to be distinguished. Fingertip-grasp means, that only the fingertips are touching the object. It is also known as precision grasp. Details can be found in [34] and chapter 4.2.1.
(e) The maximum force can not be exerted continuously. At a periodic force excitation of 25 % of the maximum force for 15 s with a pause of additional 15 s results in an impression of fatigue after 10 min. At 15 % of the maximum force this impression is reached not until 104 min [286].
(f) With the key-grasp the thumb touches the lateral part of the index finger.
(g) The test subjects were asked to exert a constant force. A monitor was used to give a visual feedback of the actual force [239].
(h) This medium error was identified to be quite independent of the exerted force between 0.25 and 1.5 N. The force had been kept constant for 14 s [239].
(i) Without visual feedback the error increased significantly and was dependant of the exerted force [239].
(j) The values were calculated of finger-forces and -geometries[87].

3.3 Conclusions from the Biology of Haptics

Next to studying the pure characteristic values of haptic perception we should keep an eye on the real meaning of μm-elongations and frequencies of 1 kHz and more and on its impact on real technical systems which is a small “finger exercise” for you to be prepared for the challenges of the design of haptic systems; the idea of this is based on a talk given by NIEMEYER at the Eurohaptics conference in 2006.

3.3.1 Stiffnesses

Already the initial touch of a material gives us information about its haptic properties. A human is able to immediately discriminate, whether he or she is touching a wooden table, a piece of rubber or a concrete wall with his or her finger tip. Besides the acoustic and thermal properties, especially the tactile and kinaesthetic feedback plays a large role. Based on the simplified assumption of a double-sided fixed plate its stiffness k can be identified by the usage of the E-modulus according to equation [158]

$$k = 2\frac{b\,h^3}{l^3} \cdot E \tag{3.8}$$

Figure 3.16a shows the calculation of stiffnesses for a plate of an edge length of 1 m and a thickness of 40 mm of different materials. In comparison, the stiffnesses of commercially available haptic systems are given in (fig. 3.16b). It is obvious that these stiffnesses of haptic devices are factors of ten lower than the stiffnesses of concrete, every-day objects like tables and walls. However, stiffness is just one criterion for the design of a good, haptic system and should not be overestimated. The comparison above shall make us aware of the fact that a pure reproduction of solid objects can hardly be realized with a single technical system. It rather takes a combination of stiff and dynamic hardware, for especially the dynamic interaction in high frequency areas dominates the quality of haptics, which has extensively been discussed in the last section.

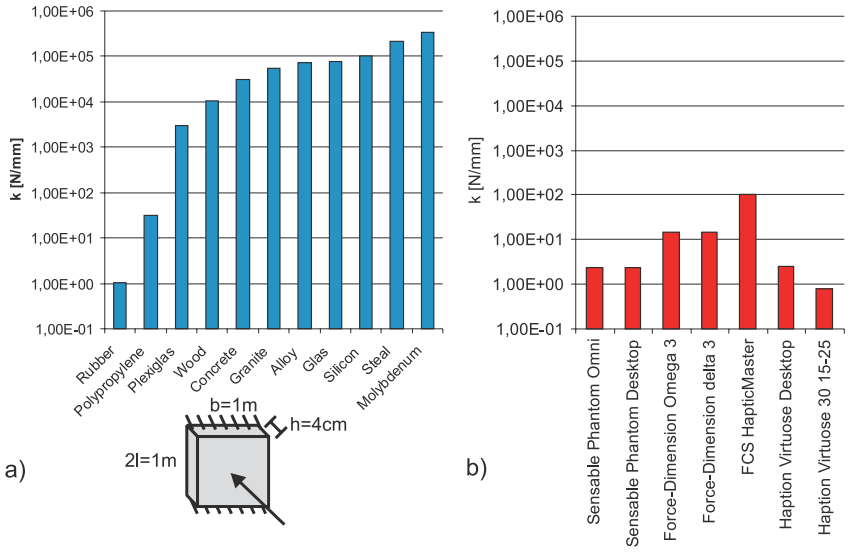


Fig. 3.16 Comparision between stiffnesses of a 1x1x0.04 m^3 plate of different materials (a) and realizable stiffnesses by commercial haptic systems (b).

3.3.2 One Kilohertz - Significance for the Mechanical Design?

Haptic perception ranges to a frequency of 10 kHz, whereby the area of highest sensitivity lies between 100 Hz and 1 kHz. This wide range of haptic perception enables us to perceive microstructures on surfaces with the same accuracy as enabling us to identify the point of impact when drumming with our fingers on a table. For a rough calculation a model according to figure 3.17 is considered to be a parallel circuit between a mass m and a spring k. Assuming an identical "virtual" volume V of material and taking the individual density ρ for a qualitative comparison, the border frequency for a step response can be calculated according to

$$f_b = \frac{1}{2\pi}\sqrt{\frac{k}{m}} = \frac{1}{2\pi}\sqrt{\frac{k}{V\rho}}. \tag{3.9}$$

Figure 3.17 shows the border frequencies of a selection of materials. Only in case of rubber and soft plastics border frequencies of below 100 Hz appear. Harder plastic material (Plexiglas) and all other materials show border frequencies above 700 Hz. One obvious interpretation would state that any qualitatively good simulation of such a collision demands at least such bandwidth of dynamics within the signal conditioning elements and the mechanical system.

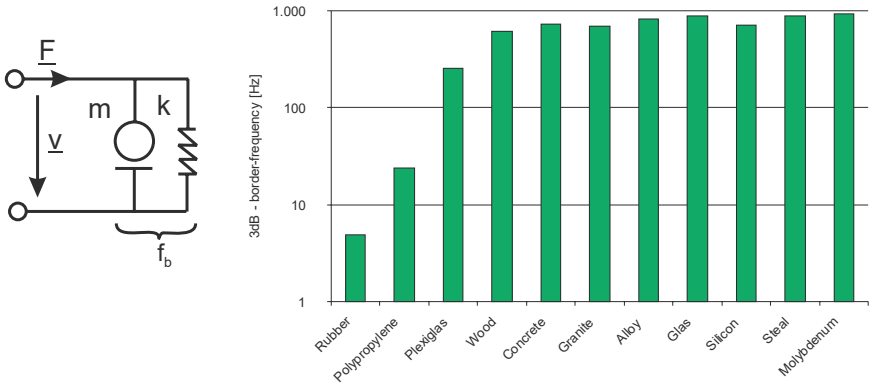


Fig. 3.17 3 dB border frequency f_b of an excitation of a simple mechanical model parametrized as different materials.

As a consequence, a frequent recommendation for the design of haptic systems is the transmission of a full bandwidth of 1 kHz (and in some sources even up to 10 kHz). This requirement is valid with respect to software and communications-engineering, as sampling-systems and algorithmic can achieve such frequencies easily today. Considering the mechanical part of the design, we see that dynamics of

1 kHz are enormous, maybe even utopian. Figure 3.18 gives another rough calculation of oscillating force amplitude according to

$$F_0 = |\underline{x} \cdot (2\pi f)^2 m|. \tag{3.10}$$

The basis of the analysis is a force source generating an output force $\underline{F}_0$. The load of this system is a mass (e.g. a knob) of 10 grams (!!). The system does not have any additional load, i.e.it does not have to generate any haptically active force to a user. A periodic oscillation of a frequency f and an amplitude $\underline{x}$ is assumed. With expected amplitudes for the oscillation of 1 mm at 10 Hz a force of approximately 10 mN is necessary. At a frequency of 100 Hz there is already a force of 2-3 N needed. At a frequency of 700 Hz the force already increases to 100 N - and this is what happens when moving a mass of 10 grams. Of course in combination with a user-impedance as load the amplitude of the oscillation will decrease in areas of below 100μm, proportionally decreasing the necessary force. But this calculation should make aware of the simple fact that the energetic design and power management of electromechanical systems with application in the area of haptics needs to be done very carefully.

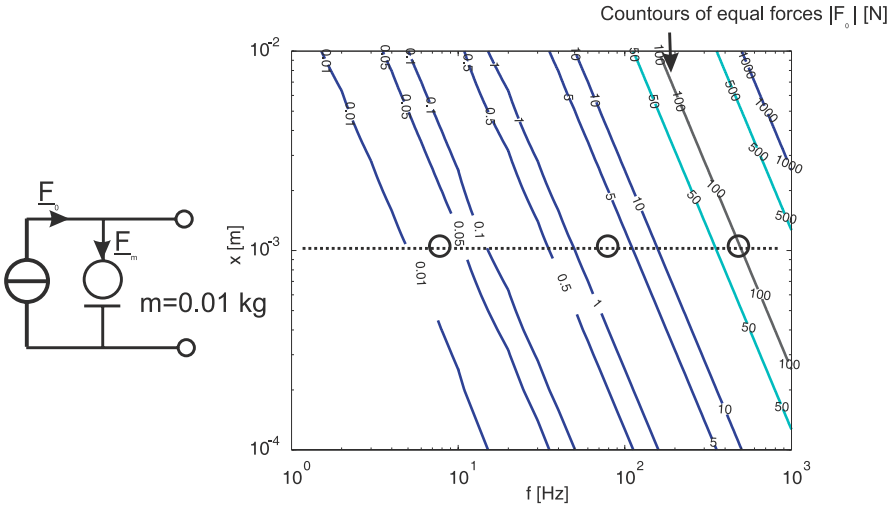


Fig. 3.18 Equipotential line of necessary forces in dependency of amplitude and frequency of the acceleration of a mass with 10 g.

The design of a technical haptic system is always a compromise between bandwidth, stiffness, dynamics of signal conditioning and maximum force-amplitudes. Even with simple systems the design process leads the engineer to the borders of what is physically possible. Therefore it is necessary to have a good model for the

user according to his being a load to the mechanical system and according to his or her haptic perception. This model enables the engineer to carry out an optimized design of the technical system and its generation is the focus point of the following chapter.

Chapter 4
Modeling the User

THORSTEN A. KERN

In the preceding chapter the frequency-dependency of mechanoreceptors relevant for haptic perception. has been discussed, i.e. the amplitude of a mechanical oscillation is perceived with varying intensity for different frequencies. Additionally, it has been shown that the dynamics of haptic perception results in demanding requirements for the design of technical systems. Consequently, a good mechanical design has to consider the user in his or her mechanical properties.

The first part of the fourth chapter deals with the discussion of the user as a mechanical *load* on the haptic device. The corresponding model is split into two independent elements depending on the frequency range of the oscillation. In the second part of the chapter a model of frequency dependent perception for high-dynamic oscillations is presented. The chapter ends with examples of the models' applications as well as a summary.

4.1 Mapping of Frequency Ranges onto the User's Mechanical Model

The frequency range of haptic perception (fig. 4.1) ranges from quasi - static (at least 10 s) to until 1 kHz and more. The area of highest sensitivity is located within the range of 100 to 300 Hz.

T.A. Kern (ed.), *Engineering Haptic Devices*, DOI 10.1007/978-3-540-88248-0_4,
© Springer-Verlag Berlin Heidelberg 2009

In fact haptic perception is far from being static. The lower border frequency however is not exactly known. A simple, personal experiment can easily be done after waking up, e.g. in bed or on a couch. Before moving any part of the body try to be aware of the position of your limbs and the structure of the cover and cushions. You will notice that after some time of rest the haptic sense will hardly transmit any information (besides thermal transfer) about your surroundings. As soon as you begin to move, all information will become immediately available. This experience is a clear hint that haptic perception does not work statically but requires at least some dynamic movement.

The area of active haptic interaction - movements, made in a conscious and controlled way by the user - is of limited range. Sources concerning the dynamics of human movements differ (subsection 3.2.3). The fastest conscious movement performed by humans is done with their fingers. Movements for typing of up to 8 Hz can be observed[1]. As these values refer to a ten-finger interaction, they have to be modified a bit. However, as the border frequency of a movement lies above the pure number of a repetitive event, an assumption of the upper border frequency of 10 Hz for active, controlled movement covers most cases. But please note that this border is not sharp and depends on the task done. The major part of the spectrum of haptic perception is passive (*passive haptic perception* and even *passive haptic interaction*). The user does not have any active influence or feedback within this passive frequency range. In fact, the user is able to modify his properties as a mechanical load by e.g. altering the force when holding a knob. But although this change influences the higher frequency range, the change itself happens with lower dynamics within the dynamic range of active haptic interaction.

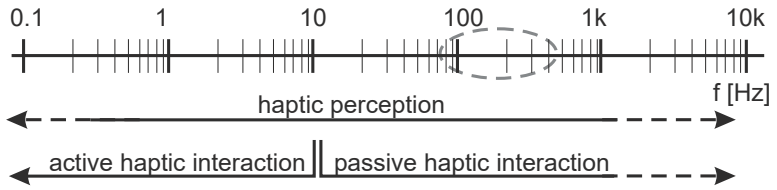


Fig. 4.1 Border-frequencies f_b resulting from an excitation of a simple mechanical model for different materials.

A look at both large classes of haptic systems (tactile and kinaesthetic systems) shows that the above modelling has slightly different impacts:

[1] 8 Hz corresponds to a typing speed of 480 keystrokes per minute. 400 keystrokes are regarded as very good for a professional typist, 300-200 keystrokes are good, 100 keystrokes can be achieved by most laymen

- The output values of kinaesthetic systems $\underline{F}_{out}$ (fig. 4.2a) result in two reactions by the user. First, a spontaneous, not directly controllable movement reaction $\underline{v}_{spo}$ happens as a result of the mechanical properties of the finger tip (depending on the type of grasp, this can be also the complete interior hand and its skin elasticity). Second an additional perception of forces takes place. This perception K [2] is weighted according to the actual situation and results in a conscious reaction of the motor parts of the body. These induced reactions $\underline{v}_{ind}$ summed up with the spontaneous reactions result in the combined output value $\underline{v}_{out}$ of the user.
- The movements of tactile devices $\underline{v}_{out}$ (fig. 4.2b) and the consciously performed movement of the user $\underline{v}_{ind}$ result in a combined movement and velocity. This elongation acts on the skin, generating the output value $\underline{F}_{out}$ as a result of its mechanical properties. This conscious movement $\underline{v}_{ind}$ sums up to $\underline{v}_{out}$ in the opposite direction of the original movement, as with opposite movement directions the skin's elongation increases and results in a larger force between user and technical system. Analogously it subtracts with movements in the same direction, as in this case the device (or the user, depending on the point of view) evades the acting force trying to keep deformation low and to perceive just a small haptic feedback. According to this model only the output value $\underline{F}_{out}$ of the combined movement is perceived and contributes to a willingly induced movement.

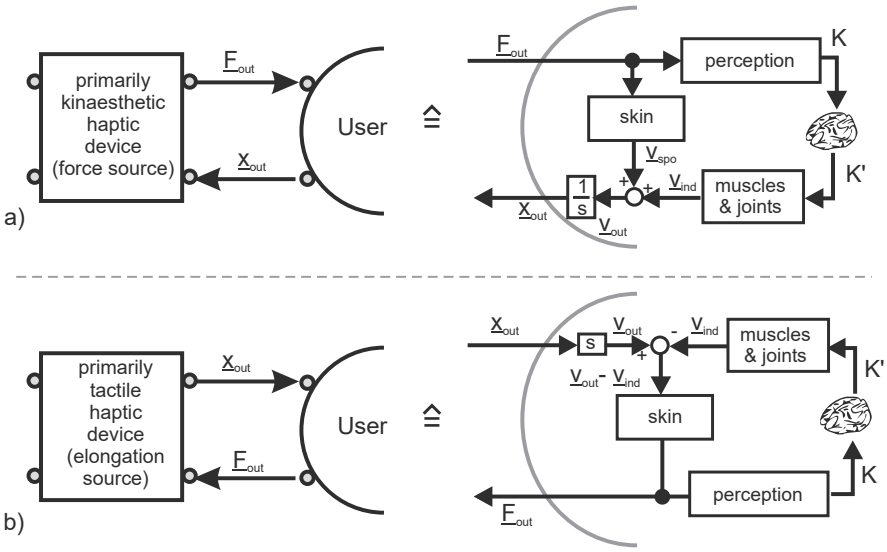


Fig. 4.2 User-models as a block-structure from kinaesthetic (a+c) and tactile (b+d) systems.

[2] K, a variable chosen completely arbitrarily, is a helpful construct for the understanding of block-diagrams rather than having a real neurological analogy

If you transfer the model of figure 4.2 into an abstract notation, all blocks correspond to the transfer-function $\underline{G}_{\mathrm{Hn}}$. Additionally, it has to be considered that the user's reaction K' is a combined reaction of complex habits and the perception K; therefore a necessity to simplify this branch of the model becomes eminent. For the purpose of device design and requirement specification, the conscious reaction is modelled by a disturbing variable only limited in bandwidth, resulting in a block-diagram according to figure 4.3c for kinaesthetic and according to figure 4.3d for tactile devices. The transfer function $\underline{G}_{\mathrm{H3}}$ corresponds to the mechanical admittance of the grasp above the border frequency of user interaction f_g. The transfer function $\underline{G}_{\mathrm{H1}}$ is a model of the quantification of haptic perception and will be discussed comprehensively in section 4.3.

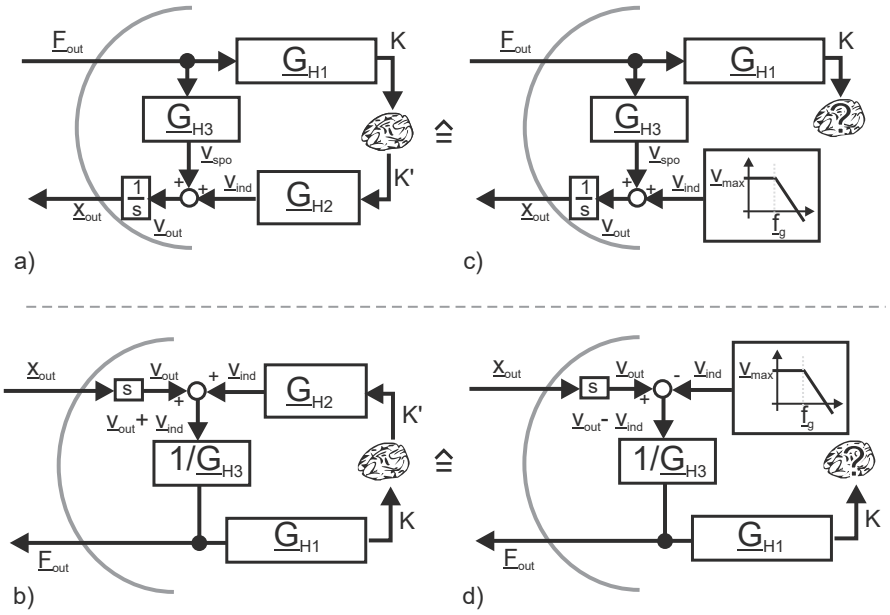


Fig. 4.3 Transformation of the user-models' block-structures in transfer-functions including simplifications of the model for the area of active haptic interaction for kinaesthetic (a+c) and tactile (b+d) systems.

With regard to the application of the presented models there are two necessary remarks to be considered:

- The notation in figures 4.2 and 4.3 for elongations $\underline{x}$ and forces $\underline{F}$ being input and output values of users is just <u>one</u> approach to the description. In fact an *impedance coupling* exists between user and haptic system making it impossible to distinguish between input and output-values. However, the decoupled haptic device is designed for being a position or force source. This in fact is the major

motivation to define input- and/or output values of the user. But there are certain actuators (e.g. ultrasonic devices) which can hardly be defined as being part of either one of these classes. As a consequence, when describing either system, the choice of the leading sign and the direction of arrows should carefully be done!

- The major motivation for this model is the description of a mechanical load for the optimized dimensioning of a haptic system. For guaranteeing the closed-loop control engineering stability of a simulation or a telemanipulation system, further care has to be taken of the frequency range of active haptic interaction below 10 Hz. Stability analysis in this area can either be achieved by more detailed models or by an observation of in- and output values according to their *control-engineering passivity*. Further information on this topic can be found in chapter 7.

The following sections on user impedance (section 4.2) and perception (section 4.3) give a concrete model for the transfer function $\underline{G}_{\mathrm{H3}}$ of figure 4.3.

4.2 A Model of the User as a Mechanical Load

The user's reaction as part of any haptic interaction combines a conscious, bandwidth-limited portion - the area of active haptic interaction - and a passive portion, mainly resulting from the mechanical properties of fingers, skin and bones. The influence of this second part stretches across the whole frequency range, but emphasizes the upper area for high frequencies. This section describes the passive part of haptic interaction. The transfer function $\underline{G}_{\mathrm{H3}}$ of figure 4.3 is a component of the impedance coupling with force-input and velocity-output and is therefore a mechanical admittance of the human $\underline{Y}_H$ (chapter 16) respectively in its reciprocal value the mechanical impedance $\underline{Z}_H$.

$$\underline{G}_{\mathrm{H3}} = \frac{\underline{v}_{\mathrm{spo}}}{\underline{F}_{\mathrm{out}}} = \frac{\underline{v}_{\mathrm{out}} - \underline{v}_{\mathrm{ind}}}{\underline{F}_{\mathrm{out}}} = \underline{Y}_H = \frac{1}{\underline{Z}_H} \tag{4.1}$$

In the following this mechanical impedance of the user will be specified. The discussion based on mechanical impedances instead of admittances allows a better approach to further understanding. An impedance is nothing else but the stiffness or hardness of a mechanics or a material. High impedance means that an object has at least one of three properties:

1. hard and stiff in the meaning of spring-stiffness,
2. inertial in the sense of inertial force,
3. sticky and tight in the sense of high friction.

In any case a small movement (velocity $\underline{v}$) results in a high force reaction $\underline{F}$ with high impedances. Low impedance means that the object, the mechanics, is accordingly soft and light. Even high velocities result in small counter forces in this case.

It is legitimate to replace the term impedance *with the term* stiffness *while reading, but one should be aware that stiffness (in the linear situation spring stiffness with a spring constant k) is only one of three characteristic properties of impedance (chapter 16).*

The human's mechanical impedance is dependent on a number of influence parameters:

- type of grasp being directly influenced by the construction of the handle
- physiological condition
- grasping force being directly influenced by the will of the user
- skin surface changes, e.g. by sweat

The quantification of human's mechanical impedance requires taking as many aspects into account as possible. The type of grasp is defined by the mechanical design of the device. Nevertheless a selection of typical grasping situations will give a good overview of typical impedances appearing during human-machine interaction. The user-individual parameters like physiological condition and skin structure can be covered best by the analysis of a large number of people of different conditions. By choosing this approach a spanwidth of percentiles can be acquired covering the mechanical impedances typically appearing with human users. The "free will" itself, however, is - similar to the area of active haptic interaction - hard if not impossible to be modelled. The time dependent and unpredictable user impedance dependency on the will can only be compensated if the system is designed to cover all possible impedance couplings of actively influenced touch. Another approach would be to indirectly measure the will to adapt the impedance model of the user within the control loop. Such an indirect measure is, in many typical grasping situations, the force applied between two fingers or even the whole hand holding an object or a handle. In the simplest design the acquisition of such a force can be done by a so called *dead-man-switch*, which in 1988 was already proposed by HANNAFORD [84] for the usage in haptic systems. A dead-man-switch is pressed as long, as the user holds the control handle in his or her hand. It detects the release of the handle resulting in a change in impedance from $\underline{Z}_H$ to 0.

4.2.1 Types of Grasps

There is a nomenclature for different types of grasps (fig. 4.4). The hand is an extremity with 27 bones and 33 muscles. It combines 13 (fingers) respectively 15

(incl. the wrist) degrees-of-freedom[3]. Accordingly the capabilities of man to grasp are extremely versatile.

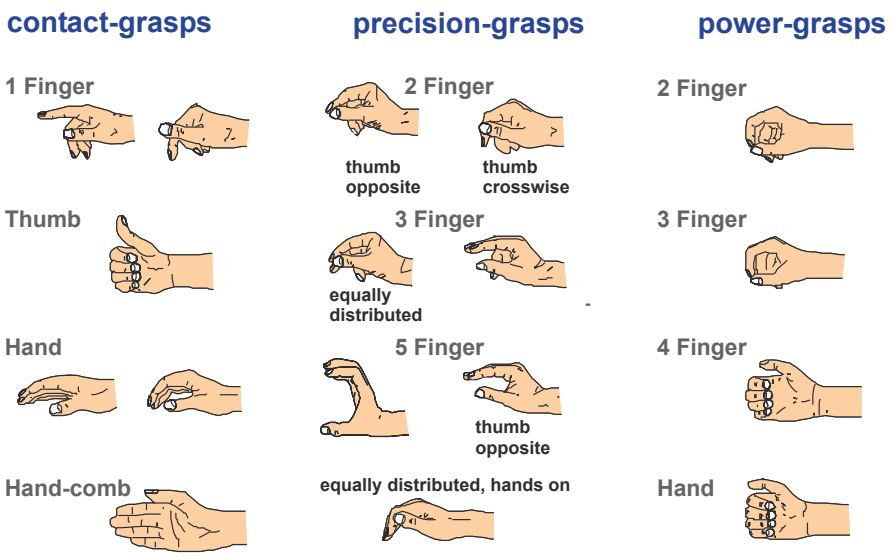


Fig. 4.4 Grasps.

There are three classes of grasps to be distinguished:

- The **contact grasp** describes the touch of an object using the whole hand or major parts of it. Keys and buttons are typically actuated by contact grasps. Even the fingers resting on a keyboard or a piano are called contact grasps. A contact grasp always blocks one direction of movement for an object (which is one half of a degree of freedom). Contact grasps can be regarded as linear only in case of a pre-load high enough. With light touches the point of release and the according lift-off of the object is always nonlinear.
- The **precision grasp** describes the grasping with several fingers. Typically a precision grasp locks at least one degree of freedom of the grasped object by form closure with one finger and a counter bearing - often another finger. Additional degrees of freedom are hindered frictionally. Precision grasps vary much in stiffness of coupling between man and machine. At the same time they are the most frequent type of grasping.
- The **power grasp** describes an object with at least one finger and a counter bearing, which may be another finger, but frequently is the whole hand. The power

[3] thumb: 4 DOF, index finger: 3 DOF, middle finger: 2 DOF (sometimes 3 DOF), ring finger: 2 DOF, small finger: 2 DOF, wrist: 2 DOF. The rotation of the whole hand happens in the forearm and therefore does not count among the degrees of freedom of the hand itself

grasp aims at locking the grasped object in all degrees of freedom by a combination of form and force closures. Power grasps are - as the name already implies - the stiffest coupling between humans and machines.

For all classes of grasps, measurements of the human's impedance can be performed. According to the approach presented by KERN [130], the measurement method and the models of user impedance are presented including the corresponding model parameters.

4.2.2 Measurement Setup and Equipment

The acquisition of mechanical impedances is a well-known, so-to-speak day-to-day problem in measurement technology. The principle of measurement is based on an excitation of the system to be measured by an actuator, simultaneously measuring force and velocity responses of the system. For this purpose combined force and acceleration sensors (e.g. the impedance sensor 8001 from Brüel & Kjær) exist, whereby the charge amplifier of the acceleration sensor includes an integrator. In general the impedance of organic systems is nonlinear and time-variant. This nonlinearity is a result of a general viscoelastic behaviour of tissue resulting from a combined response of relaxation, conditioning, stretching and creeping [63]. These effects can be reproduced by mechanical models with concentrated elements. However, they are dependent on the time-history of excitation to the measured object. It can be expected that measurements based on step excitation are different from those acquired with a sinoid sweep. Additionally, the absolute time for measurement has some influence on the measures by conditioning. Both effects are systematic measurement errors. Consequently, the models resulting from such measurements are an indication of the technical design process and should always be interpreted with awareness of their variance and errors.

All impedance measures presented here are based on a sinoid-sweep from upper to lower frequencies. The excitation has been made with a defined force of 2 N amplitude at the sensor. The mechanical impedance of the handle has been measured by calibration measurements and was subtracted from the measured values. The impedance-sensors are limited concerning their dynamic and amplitude resolution, of course. As a consequence, the maximum frequency up to which a model is valid depends on the type of grasp and its handle used during measure. This limitation is a direct result of the amplitude resolution of the sensors and the necessity at high frequencies to have a significant difference between the user's impedance and the handle's impedance for the model to be built on. The presented model-parameters are limited to the acquired frequency range and cannot be applied to lower or higher frequencies. The measurement setup is given in figure 4.6.

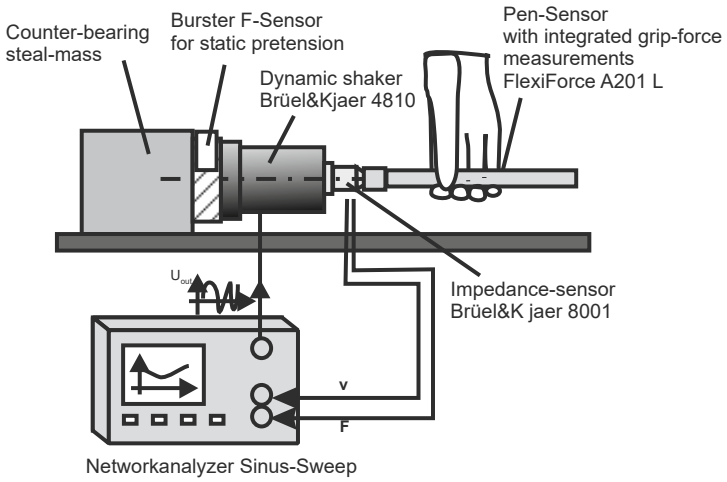


Fig. 4.5 Measurement setup for the acquisition of user-impedances according to [130].

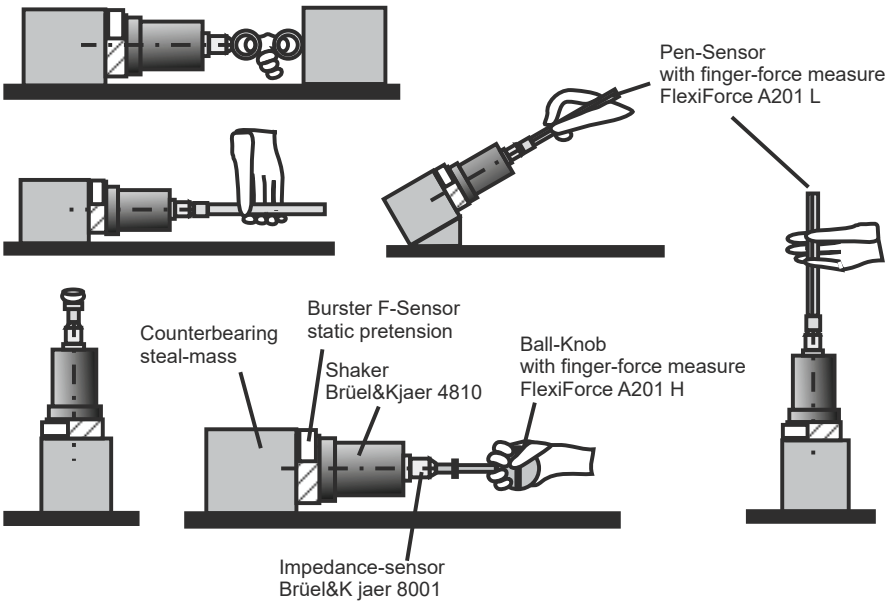


Fig. 4.6 Impedance measurement settings for different grasps.

4.2.3 Models

In order to approximate the human's impedance a number of different approaches were taken in the past (fig. 4.7). For its description mechanical models based on concentrated linear elements were chosen. They range from models including active user reactions represented by force sources (fig. 4.7a), to models with just three elements (fig. 4.7c) and combined models of different design. The advantage of a mechanical model compared to a defined transfer function with a certain degree in enumerator and denominator results from the possibility of interpreting the elements of the model as being a picture of physical reality. Elasticities and dampers connected in circuit with the exciting force can be interpreted as the coupling to the skin. Additionally the mechanical model creates very high rankings by its interconnected elements which allow a much better fit to measurements than free transfer functions.

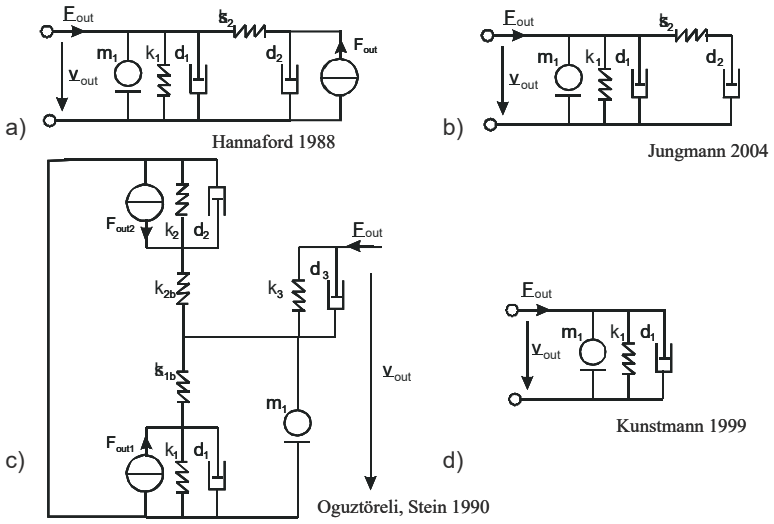


Fig. 4.7 Modelling the user with concentrated elements, (a) [84], (b) [119] (c) [188], (d) [147].

KERN [130] defined an eight-element model based on the models in figure 4.7 for the interpolation of the performed impedance measures. The model can be characterized by three impedance groups typical for many grasping situations:

$\underline{Z}_3$ (equ. 4.4) models the elasticity and damping of the skin being in direct contact with the handle. $\underline{Z}_1$ (equ. 4.2) is the central element of the model and describes the mechanical properties of the dominating body parts - frequently fingers. $\underline{Z}_2$ (equ. 4.3) gives an insight into the mechanical properties of the limbs, frequently hands, and allows to make assumptions about the pre-loads in the joints in a certain grasping situation.

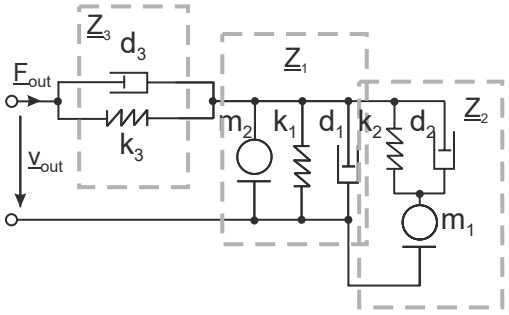


Fig. 4.8 Eight-element model of the user's impedance [130], modeling the passive mechanics for frequencies >20 Hz.

$$\underline{Z}_1 = \frac{s^2 m_2 + k_1 + d_1 s}{s} \tag{4.2}$$

$$\underline{Z}_2 = \left(\frac{s}{d_2 s + k_2} + \frac{1}{s m_1} \right)^{-1} \tag{4.3}$$

$$\underline{Z}_3 = \frac{d_3 s + k_3}{s} \tag{4.4}$$

$$\underline{Z}_B = Z_1 + Z_2 \tag{4.5}$$

Combined, the model's transformation is given as

$$\underline{Z}_\mathrm{H} = \underline{Z}_3 \| \underline{Z}_B \tag{4.6}$$

$$\underline{Z}_\mathrm{H} = \left(\frac{s}{d_3 s + k_3} + \left(\frac{s^2 m_2 + k_1 + d_1 s}{s} + \left(\frac{s}{d_2 s + k_2} + \frac{1}{s m_1} \right)^{-1} \right)^{-1} \right)^{-1} \tag{4.7}$$

4.2.4 Modeling Parameters

For above model (equ. 4.7) the mechanical parameters can be identified by measurement and approximations with real values. For the values presented here approximetely 48 to 194 measurements were made. The automated algorithm combines an evolutionary approximation procedure followed by a curve-fit with optimisation based on NEWTON [30] curve fitting, to achieve a final adjustment of the evolutionarily found starting parameters according to the measurement data. The measure-

ments vary according to the mechanical pre-load - the grasping force - to hold and move the control handles. This mechanical pre-load was measured by force sensors integrated into the handles. For each measurement this pre-load could be regarded as being static and was kept by the subjects with a 5% range of the nominal value. As a result the model's parameters could be quantified not only dependent on the grasping situation but also dependent on the grasping force. The results are given in the following section. The display of the mechanical impedance is given in decibel, whereby 6 dB equals a doubling of impedance. The list of model values for each grasping situation is given in appendix chapter 18.

4.2.4.1 Power Grasps

Within the class of power grasps three grasping-types were analyzed. Impedance between 35 and 45 dB is measured for the grasp of a cylinder (fig. 4.9) and of a sphere of similar dimensions (fig. 4.10) with the whole hand. It shows an anti-resonance in the area of 80 Hz, which moves for a grasp of the cylinder slightly to higher frequencies with increased grasping force. The percentiles, especially the 5th, reveal that the model is based on a large variance and uncertainty. It is likely that the influence of the subjects' variability in their physical parameters like the size of hands and fingers influence these measurements a great deal. In case of grasping two rings with thumb and index finger (fig. 4.11) the results are much more accurate. Impedance ranges in between 15 and 35 dB. The anti-resonance in the frequency range of 70 to 100 Hz shows a clear dependency on grasping forces. If we look at the parameters, this change is a result of a variance within the elasticity coefficients k_1 and m_2, building the central parallel resonance of the model. The mechanical system "hand" becomes stiffer (k_1 increases), but the mass m_2 part of the anti-resonance diminishes. An easy interpretation of this effect is not obvious. At a value of 10 kg the mass m_1 builds an almost stiff counter bearing.

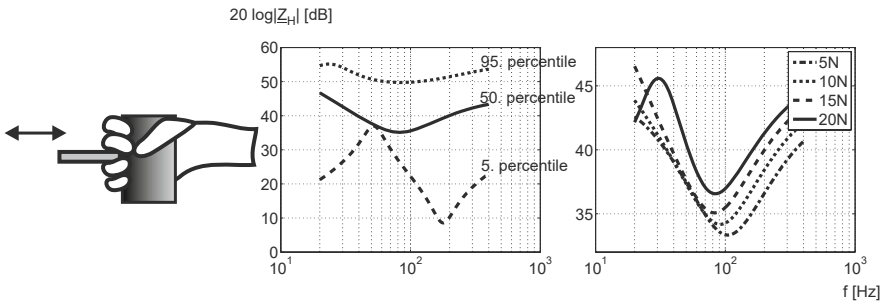


Fig. 4.9 Impedance with percentiles (a) and at different force levels (b) for power grasps of a cylinder (Ø25 mm, defined for 20 Hz to 400 Hz).

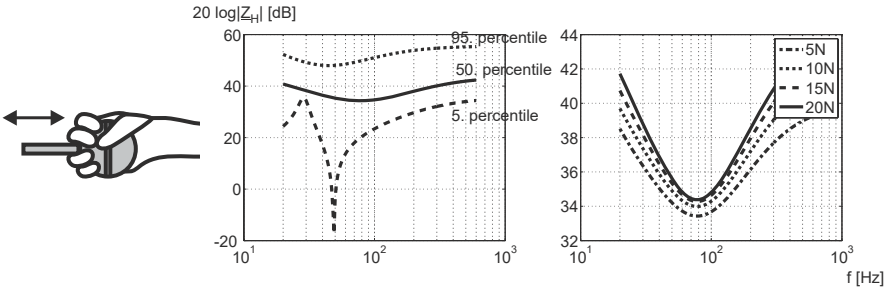


Fig. 4.10 Impedance with percentiles (a) and at different force levels (b) for a power-grasp of a sphere (Ø40 mm, defined for 20 Hz to 600 Hz).

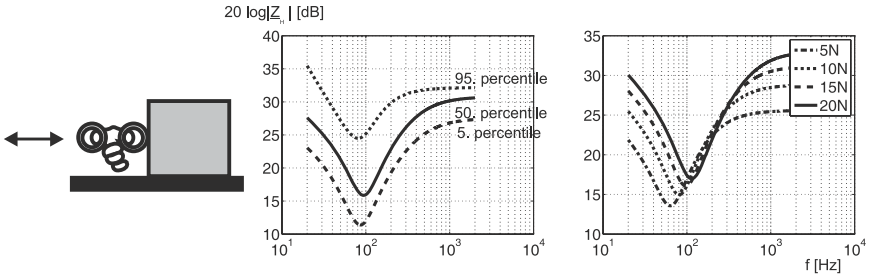


Fig. 4.11 Impedance with percentiles (a) and at different force levels (b) for a two-fingered power-grasp of rings (Ø25 mm of the inner ring, defined for 20 Hz to 2 kHz).

4.2.4.2 Precision Grasps

Within the area of precision grasps three types of grasps were analyzed. Holding a measurement cylinder similar to a normal pen in an angle of 30° (fig. 4.12), we find a weak anti-resonance in the area of around 150 to 300 Hz. This anti-resonance is dependent on the grasping force and moves from weak forces and high frequencies to large forces and lower frequencies. The general dependency makes sense, as the overall system becomes stiffer (the impedance increases) and the coupling between skin and cylinder becomes more efficient resulting in more masses being moved at higher grasping forces. The general impedance does not change significantly. if the cylinder is held in a position similar to a máobi Chinese pen (fig. 4.13). However the dependency on the anti-resonance slightly diminishes compared to the above pen hold posture. This is completely different to the variant of a pen in a horizontal position held by a three finger grasp (fig. 4.14). A clear antiresonance with frequencies between 80 and 150 Hz appears largely dependent in shape and position on the grasping force. All observable effects in precision grasps can hardly be traced back to the change of a single parameter but are always a combination of many parameters' changes.

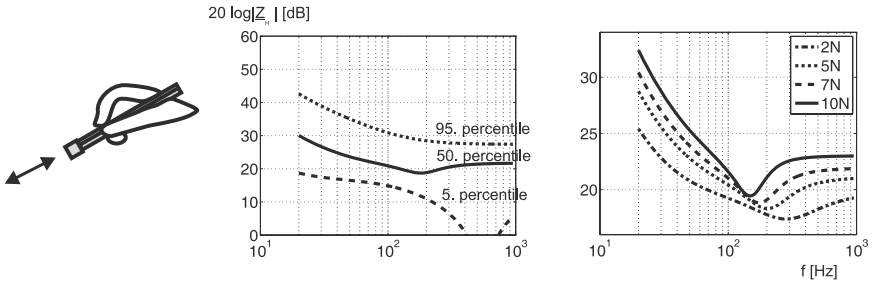


Fig. 4.12 Impedance with percentiles (a) and at different force levels (b) for a two fingered precision-grasp of a pen-like object held like a pen (Ø10 mm, defined for 20 Hz to 950 Hz).

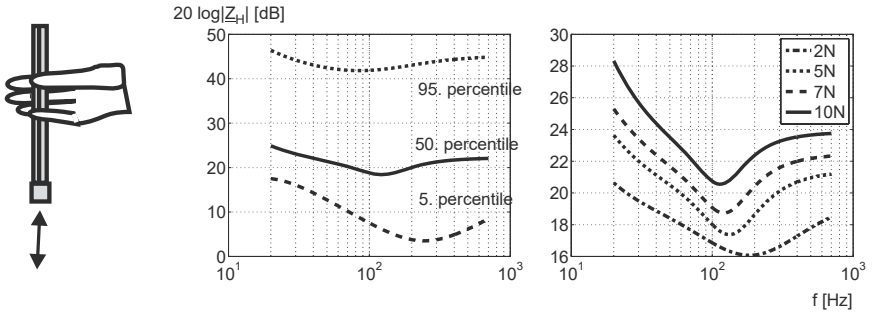


Fig. 4.13 Impedance with percentiles (a) and at different force levels (b) for a two fingered precision-grasp of a pen-like object held like a "máobi" Chinese pen (Ø10 mm, defined for 20 Hz to 700 Hz).

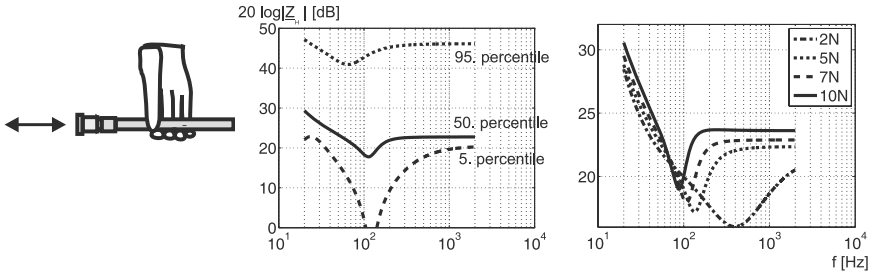


Fig. 4.14 Impedance with percentiles (a) and at different force levels (b) for a five fingered precision-grasp of a pen-like object in horizontal position (Ø10 mm, defined for 20 Hz to 2 kHz).

4.2.4.3 One-Finger Contact Grasp

All measurements were done on the index finger. Parts of these results were published in EHC 2008 [132] and are repeated for completeness here. Direction of touch, size of touched object and touch-force normal to the skin were varied within this analysis. Figure 4.15a shows the overview of the results for a touch being analyzed in normal direction. The mean impedance varies between 10 and 20 dB with a resonance in the range of 100 Hz. Throughout all measured diameters of contactor size and forces, no significant dependency of the position of the anti-resonance on touch forces were noted. However, a global increase in impedance is clearly visible. Observing the impedance dependent on contactor size, we can recognize an increase of the anti-resonance frequency. Additionally, it is fascinating to see that the stiffness decreases with an increase of contact area. The increase in resonance is probably a result of less material and therefore less inertia participating in generating the impedance. The increase in stiffness may be a result of smaller pins deforming the skin more deeply and therefore getting nearer to the bone as a stiff mechanical counter bearing. In comparison, with measurements performed with a single pin of only 2 mm in diameter (fig. 4.15b), the general characteristic of the force dependency can be reproduced. Looking at the largest contact element of 15 mm,in diameter, we are aware of a movement of the resonance frequency from 150 Hz to lower values down to 80 Hz for an increase in contact force. In orthogonal direction the skin results differ slightly. Figure 4.16a shows a lateral excitation of the finger pad with an obvious increase of impedance at increased force of touch. This rise is mainly a result of an increase of damping parameters and masses. The position of the anti-resonance in frequency domain remains constant at around 150 Hz. The picture changes significantly for the impedance in distal direction (fig. 4.16b). The impedance still increases, but the resonance moves from high frequencies of around 300 Hz to lower frequencies. Damping increases too, resulting in the anti-resonance being diminished until non-existence. At 45° (fig. 4.16c) a combination of both effects appears. Anti- resonance moves to a higher frequency and loses its sharpness compared to the pure lateral excitation. A first trend of change within the position of the anti-resonance in frequency domain with higher forces can be identified additionally.

4.2.4.4 Superordinate Comparison of Grasps

It is interesting to compare the impedances among different types of touch and grasps with each other:

1. Almost all raw data and the interpolated models show a decrease of impedance within the lower frequency range of 20 Hz to the maximum of the first anti-resonance. As to precision grasps (fig. 4.12 to 4.14), normal fingertip excitation (fig. 4.15) and the power grasp of rings (fig. 4.11) the gradient equals 20 dB/decade resembling a dominating pure elongation proportional effect of

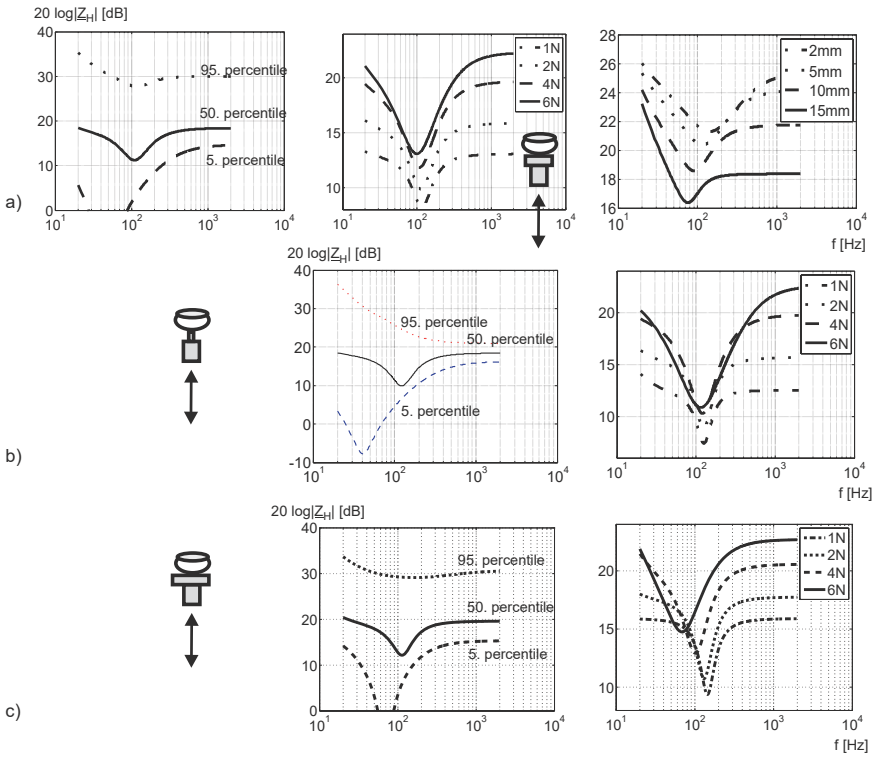


Fig. 4.15 Impedance of finger touch via a cylindrical plate for different contact forces (1 to 6 N) and in dependency from diameter (a), for the smallesplate(Ø 2 mm) and the largest plate (Ø 15 mm) (defined for 20 Hz to 2 kHz).

force response - elasticity - within a low frequency range. Within this low bandwidth-area nonlinear effects of tissue including damping seem to be not very relevant. Looking at this type of interactions we can assume that any interaction including joint rotation of a finger is almost purely elastic in a low frequency range.

2. Many models show a clear antiresonance. Its position varies between 70 Hz at power grasps (fig. 4.9 to 4.11) and 200 Hz or even 300 Hz at finger touch analyzed in orthogonal direction (fig. 4.16). The resonance is a natural effect of any system including a mass and elasticity. Therefore it is not its existence which is relevant for interpretation, but its shape and the position within the frequency range. As to positions, especially the power grasp of two rings (fig. 4.11), the precision grasps of a cylinder in a pen-like position (fig. 4.12) and in horizontal position (fig. 4.14) and the touch of an orthogonal moving plate in distal direction (fig. 4.16c) and a large plate in normal direction (fig. 4.16c) have a clear dependence on grasping force. The interpretation is not as obvious as in case

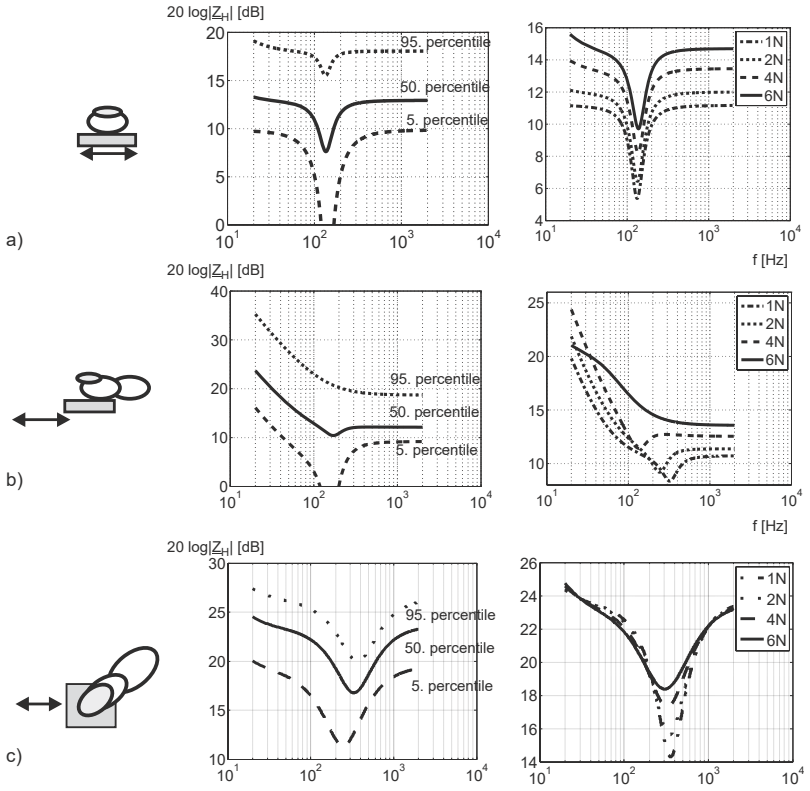


Fig. 4.16 Impedance for finger touch of a plate moving in orthogonal direction to the skin at different force levels (1 to 6 N) (defined for 20 Hz to 2 kHz). Movement in lateral direction (a), distal direction (b) and at 45° (c)..

one. We assume that the normal touch of the plate shows similarities to the contact situation when touching the rings. Additionally the normal touch is part of the precision grasps mentioned above. In the case of many subjects grasping the horizontal cylinder, it could be observed that the thumb was positioned less orthogonally but more axially to the cylinder, which could excite it primarily in distal direction, thus also contributing to this effect.

3. The shape of the anti-resonance is another interesting factor. It can be noted that especially in the analysis of finger grasps and there at orthogonal excitation (fig. 4.16a), the anti-resonance is very narrow. An interpretation is hard to be formulated. It seems that with grasps and especially touches involving less material the anti-resonance becomes narrower in shape.
4. For all measurements, at high frequencies above the anti-resonance, the frequency characteristic becomes linear and constant, which resembles a pure damping effect. This becomes obvious at the pen-hold posture among the preci-

sion grasps (fig. 4.12) and with the lateral displacement in orthogonal direction, (fig. 4.16a), but is part of any curve and model. Alternatively, inertia could be assumed to dominate the high frequencies, being represented by a linear increase of mechanical impedance. Mainly power grasps show a tendency to this increase. This measured effect is especially relevant, as it confirms common assumptions that for high frequency haptic playback with kinaesthetic devices, the user can be assumed as a damping load.

5. A last glance should be taken at the absolute height level and the variance of height of the impedance due to pre-loads. For all grasps it varies in a range (regarding the median curves only) of 20 dB as a maximum. Impedance is higher for power grasps, slightly lower for precision grasps and very much lower for touches, which is immediately obvious. The change in the pre-load for one grasp typically displaces the absolute impedance to higher levels. This displacement varies between 4 to 10 dB.

If speculations should be made on still unknown, not yet analyzed types of touches according to the given data, it should be reasonable to assume the following:

A. Power grasp:

The median impedance should be around 36 dB. Model the impedance with a dominating elasticity effect until an anti-resonance frequency of 80 Hz, not varying much neither in height nor in position of the anti-resonance. Afterwards allow inertia to dominate the model's behaviour.

B. Precision grasp:

The median impedance should be around 25 dB. Model the impedance with a dominating elasticity effect until an anti-resonance frequency of around 200 Hz. The position of the anti-resonance diminishes in an area of 100 Hz due to change in preload. Above that anti-resonance let the impedance become dominated by a damping effect. The height of impedance changes in a range of 5 dB by the force of the grasp.

C. Finger touch:

The median impedance should be around 12 dB. Model the impedance with a well balanced elasticity and damping effect until an anti-resonance frequency of around 150 Hz. The position of the anti-resonance is quite constant, with the exception of large contact areas moving in normal and in distal direction. Above that anti-resonance let the impedance become strongly dominated by a damping effect. The

absolute height of impedance changes in an area of up to 10 dB depending on the force during touch.

4.2.4.5 Comparison with Existing Models

For further insight into and qualification of the results, a comparison with published mechanical properties of grasps and touches is presented in this section. There are two independent trends of impedance analysis in the scientific focus: the measurement of mechanical impedance as a side product of psychophysical studies at threshold level, and measurements at higher impedance levels for general haptic interaction. The frequency plots of models and measurements are shown in figure 4.17. In [110] the force detection thresholds for grasping a pen in normal orientation have been analyzed. Figure 4.17a shows an extract of the results compared to the pen-like grasp of a cylinder of the model in figure 4.12a. Whereas the general level of impedance does fit, the dynamic range covered by our model is not as big as described in literature. Analyzing the data as published, we can state that the minimum force measured by ISRAR is $\approx 60\,\mu$N at the point of lowest impedance. A force sensor reliably measuring at this extreme level of sensitivity exceeds the measurement error of our setup and may be the explanation of the difference in the dynamic range covered. In another study [111] the force detection threshold of grasping a sphere with the finger tips was analyzed. The absolute force level of interaction during these measurements was in the range of millinewton. A comparison (fig. 4.17b) between our model of touching a sphere and these data show a difference in the range of 10 to 20 dB. However such small contact forces resemble a large extrapolation of our model data to low forces. The difference can therefore be easily explained by the error resulting from this extrapolation. FU [61] measured the impedance of the finger tip at a low force of 0.5 N. He advanced an approach published by HAJIAN [82]. A comparison between our model and their data concerning the shape is hardly possible due to the little number of discrete frequencies of this measurement. However the impedance is again 10 dB lower than of our touch model of a five millimeter cylinder at normal oscillations similar to figure 8a. Once more the literature data describe a level of touch force not covered by our measurements and therefore the diagram of Fig. 10c is an extrapolation of the model of these low forces. As a conclusion of this comparison, the model presented here cannot necessarily be applied to measurements done at lower force levels. Publications dealing with touch and grasp at reasonable interaction forces reach nearer to the model parameter estimated by our research. YOSHIKAWA [293] published a study of a three element mechanical model regarding the index finger. The study was based on a time-domain analysis of a mechanical impact generated by a kinaesthetic haptic device. The measured parameters result in a frequency plot (fig. 4.17d) which is comparable to our model of low frequencies, but does neither show the complexity nor the variability of our model in a high frequency range of above 100 Hz. A similar study in time-domain was performed by HAJIAN [82] with just slightly different results. Measurements

available as raw data from JUNGMANN [120] taken in 2002 come quite close to our results, although obtained with different equipment. Besides these frequency plots, the model's parameters allow a comparison with absolute values published in literature: SERINA [224] made a study on the hysteresis of the finger tips' elongation vs. force curve during tapping experiments. This study identified a k-factor for pulp stiffness ranging from 2 N/mm at a maximum tapping-force of 1 N to 7 N/mm at a tapping force of 4 N. This k-factor is in the area of 3 to 8 times larger than the dominating k-Faktor k_2 in our eight-element model. The results of FU [61] make us assume that there was a systematic error concerning the measurements of SERINA, as the elongation measured at the fingernail does not exclusively correspond to the deformation of the pulp. Therefore the difference in k-Factor between our model and their measurements can become reasonable. Last but not least MILNER [177] carried out several studies on the mechanical properties of the finger tip in different loading directions. In the relevant loading situation a k-factor of 200 N/m to 500 N/m was identified by him. This is almost perfect within the range of our model's stiffness.

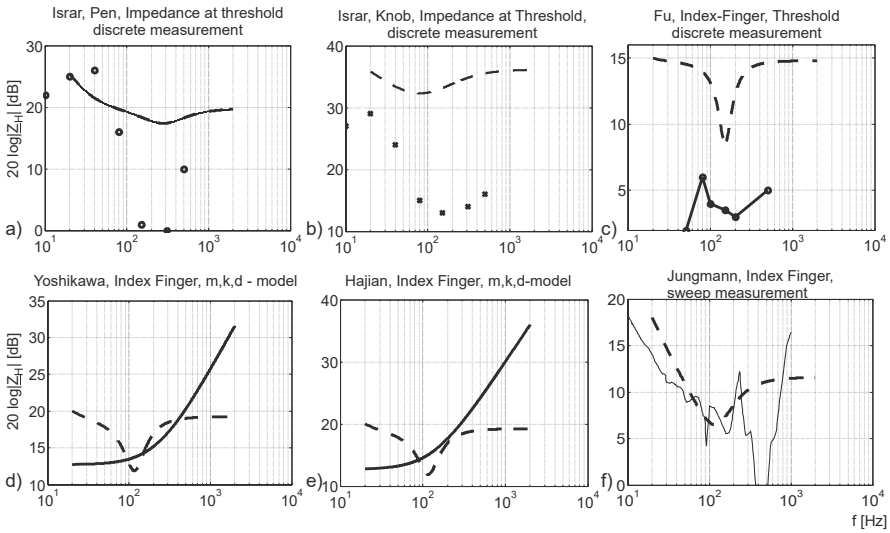


Fig. 4.17 Comparision of the model from figure 4.8 with data from similar touches and grasps as published by ISRAR [110, 111], FU [61], YOSHIKAWA [293], HAJIAN [82], JUNGMANN [120].

4.2.4.6 Final Remarks on Impedances

The impedance model as presented here will help with the modelling of haptic perception in high frequency ranges of above 20 Hz. However, it completely ignores

any mechanical properties below that frequency range. This is a direct consequence of the general approach to human machine interaction presented in section 4.1 and has to be considered when using this model. Another aspect to consider is that the above measurements show a large inter-subject variance of impedances. In extreme cases they span 20 dB meaning nothing else but a factor of 10 between e.g. the 5th and the 95th percentile. Further research on the impedance models will minimize this variance and allow a more precise picture of impedances. But already this database, although not yet completed, allows to identify helpful trends for human load and haptic devices.

4.3 Modelling Haptic Perception

At the time of this book going to press no standardized and accredited procedure for the quantification of dynamic haptic perception exists. Nevertheless there is a need to transform haptic perception into values showing at least some correlation with each other, allowing at least some comparison of sources with haptically relevant outputs. When designing a technical system it is of obvious advantage to have a tool enabling the engineer to judge the quality of the technical design. This may be achieved by time discrete simulation or via analytical identification of characteristic values. Within optics and acoustics such procedures and methods have been known for decades. By the aid of $v(\lambda)$ and $v'(\lambda)$ the brightness of a light source or an illuminated object is weighted in a spectral domain according to a normalized perception curve of wavelengths. In acoustics the spectrum of a dynamic pressure is calculated according to the perception curve in phone and sone. Surprisingly no such method has been established for the haptic sense. Most likely this is a direct consequence of the diversity of haptic interaction. The large number of possible touch situations and the changing load and perception-situations which are already considerable, are further increased by the time dependent change of the interaction itself. Only the haptic sense feeds back on the object generating the sense, something which also happens to a certain extent to the visual or acoustic sense but mostly can be ignored[4].

It has to be clearly distinguished between parameters describing the quality of a technical system, like the transparency (section 2.2.3), and a value quantifying haptic perception. The parameters of the technical system can be part of the optimization of a device's haptic performance, but they are just one part of the technical transmission.

Freeing oneself from the definitions made in different disciplines and focusing on the engineering aspect of *haptic perception*, we could regard the following definition as a possible approach:

[4] Obviously, electromagnetic radiation at visual wavelengths is perceived e.g in a crowd by a spectator absorbing photons, which will afterwards no longer be available to another spectator further behind. People cast shadows - yet this is typically trivial and not relevant for the design of a visual display. Similarly, the acoustics of a room is changed by the presence of a person in the room. This effect is considered for professional acoustic installations especially in theatres or operas as being of high importance for their performance. But in every- day-life situations and for the majority of artificial sound sources the presence of the auditor is not considered at all.

"The sensual impression of an acting force describes the perception of the presence of a mechanical energy, being not yet perceived as a position change." (T.A. Kern)

The skin, muscles and joints function like force sensors. Smallest elongations are a direct result of acting forces. These deformations are measured directly by the mechanoreceptors known from chapter 3. The skin itself, however, - in contrast to the design aims of any technical system - is highly nonlinear in its force/displacement curve and the distribution of the primary sensors.

In analogy to force sensors, the haptic perception of skin, muscles and joints has to include transfer characteristics describing the frequency dependency of the perception of the input value "force". Additionally, it should be possible to identify a mechanical load resulting from the force sensor for the system to be measured. Finally there has to be a minimum measurement value as being the result of the noise on the sensors, defining a minimum measurable force and an absolute and relative error. The existence of all these properties has already been known from haptic perception, although hidden within different terminologies. The transfer characteristics are known as a perception curve of mechanical oscillations. The borderline of minimum measurement error is known as JNDs for different contact situations. The DL and the power-law provide us with values for the absolute and relative measurement errors. The crux of the quantification of haptic perception lies in the fact that characteristic values and measurement methods have developed independently in several scientific disciplines throughout decades. They show gaps in the documentations of experiments, in the statistic data base or in the number of performed experiments and in the observation of influence parameters. Additionally, these values scatter widely, as it is the case with many psychophysical experiments, requiring a large number of experiments for a large number of touch situations. Consequently it is necessary to formulate working hypotheses taking from the number of publications which need an elaborate and long-term verification with very controlled experiments. Among others, this kind of research is done especially at two institutes[5]. The finalization and verification of the method described in the following chapters will take several more years for a concluding proof. Nevertheless it has already shown some very helpful qualities for the technical design process and for assessing the resulting haptic properties of a product. This is why it is described here. Another reason is the author's intention to encourage readers to pursue this way of thinking and to make use of its results.

As already speculated in section 4.2.4.3, it is most likely that the deformation of the skin is proportionally linked to the intensity of haptic perception. GESCHEIDER and many others have proved by measurements of the perception thresholds of mechanical oscillations (section 3.2.2), that there is a functional coherence - similar to a transfer characteristic - between physically measurable amplitude and a physio-

[5] Institute for Electromechanical Design, Technische Universität Darmstadt, Germany; Haptic Interface Research Laboratory, Purdue University, Indiana, USA

logical and/or psychological perception. Looking at the structure of skins and the orientation of receptors, we can state that deformations and elongations have to be responsible for the perception of mechanical stimuli.

As begun in section 4.1, the transfer function $\underline{G}_{H1}$ shall put the physical force $\underline{F}_{\text{out}}$ and the perception of this force $\underline{K}$ into relation with each other.

$$\underline{G}_{H1} = \frac{\underline{K}}{\underline{F}_{\text{out}}} \tag{4.8}$$

The characteristic curve known best (fig. 3.15 on page 50) merely describes the perception of mechanical oscillations. There is a transfer function according to

$$\underline{G}_{\text{FIP}} = \frac{\underline{K}}{\underline{\underline{x}}}, \tag{4.9}$$

whereby the index FIP describes the conversion factor between an impression of an oscillation and the mechanical oscillation itself, independently of whether it is an elongation x or a force-amplitude F.

Figure 3.15 in section 3.2.2 shows exactly this perception threshold for the haptic perception of mechanical oscillations at the finger in dependency of the contact area. The curves can be approximated (fig. 4.18) and be normalized on their particular maximum. For a contact area of $A = 2.9\,\text{cm}^2$ the following approximation can be applied

$$G_{\text{FIP A2.9}} = \frac{K\,(1+T_{n1}\,s)\,(1+T_{n2}\,s)^4\,(1+T_{n3}\,s)^3}{(1+T_{p1}\,s)\,(1+T_{p2}\,s)^9\,(1+T_{p3}\,s)^2\,(1+T_{p4}\,s)^2} \tag{4.10}$$

with coefficients taken from table 4.1. For a contact area of $A = 0.008\,\text{cm}^2$ an analogue approximation can be made

$$G_{\text{FIP .008}} = \frac{K\,(1+T_{n1}\,s)^2}{(1+T_{p1}\,s)^2\,(1+T_{p2}\,s)^9\,(1+T_{p3}\,s)^8} \tag{4.11}$$

with coefficients taken from table 4.2.

Within the frequency range of 10 Hz to 10 kHz the resulting curves are appropriate to weight oscillations according to haptic perception. As the user impedance models of section 4.1 are based on the perception of forces, it is necessary to introduce a conversion factor to get from equation 4.9 to 4.8.

Our own experiments and the results obtained by others [111] allow the assumption about the user's impedance $\underline{Z}_H$ as being appropriate as conversion factor. The conclusion drawn from the modelling in section 4.2 shows that the transfer function of equation 4.1 is nothing else but the user's mechanical admittance, as can be seen in

$$\underline{G}_{H3} = \frac{1}{\underline{Z}_H} \,\text{mit}\, \underline{Z}_H = \frac{\underline{F}_{\text{out}}}{\underline{v}_{\text{spo}}} \tag{4.12}$$

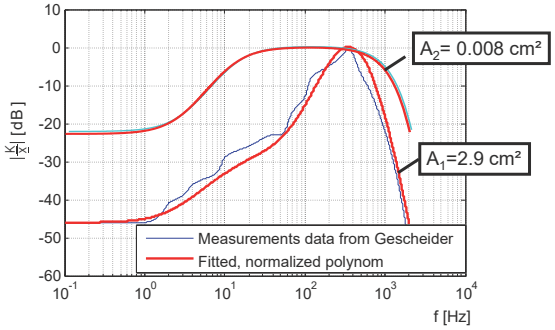


Fig. 4.18 Approximation of the characteristic curves according to GESCHEIDER from figure 3.15.

With this coherence and the knowledge that $\underline{x} = \frac{\underline{v}}{s}$, one result of equation 4.9 is

$$\underline{G}_{\mathrm{FIP}} = \frac{\underline{K}}{\underline{x}} = \frac{\underline{K}\,s}{\underline{v}_{\mathrm{spo}}} = \frac{\underline{K}\,s\,\underline{Z}_H}{\underline{F}_{\mathrm{H}}} \tag{4.13}$$

and consequently for $\underline{G}_{H1}$ (with $\underline{F}_{\mathrm{out}} = -\underline{F}_{\mathrm{H}}$) the equation is

$$\underline{G}_{H1} = \frac{\underline{K}}{\underline{F}_{\mathrm{out}}} = -\frac{\underline{G}_{\mathrm{FIP}}}{s\,\underline{Z}_H} \tag{4.14}$$

Equation 4.14 now provides the possibility to quantify the perceived quality of a technological haptic system with an output force based on assumptions on user impedance $\underline{Z}_H$ and the characteristic curve of haptic perception $\underline{G}_{\mathrm{FIP}}$.

Table 4.1 Parameters for the approximation in figure 4.18 for a contact area of $2,9\,\mathrm{cm}^2$ according to equation 4.10.

Parameter	**Value**
K	$5\,10^{-3}$
T_{n1}	$(2\,\pi\,2)^{-1}$
T_{n2}	$(2\,\pi\,80)^{-1}$
T_{n3}	$(2\,\pi\,320)^{-1}$
T_{p1}	$(2\,\pi\,15)^{-1}$
T_{p2}	$(2\,\pi\,200)^{-1}$
T_{p3}	$(2\,\pi\,420)^{-1}$
T_{p4}	$(2\,\pi\,1000)^{-1}$

Table 4.2 Parameters for the approximation in figure 4.18 for a contact area of $0.008\,\mathrm{cm}^2$ according to equation 4.11.

Parameter	Value
K	0.0785
T_{n1}	$(2\pi 3)^{-1}$
T_{p1}	$(2\pi 11)^{-1}$
T_{p2}	$(2\pi 3000)^{-1}$
T_{p3}	$(2\pi 4000)^{-1}$

4.4 Application Remarks

The model in figure 4.3 can be transferred to a mechanical circuit based on concentrated elements (fig. 4.19). This simplifies the formulation of interaction-interdependencies between the human and the technological system, and allows the analysis of some simple examples.

4.4.1 Kinaesthetic System

For a kinaesthetic system (fig. 4.19) a simplified model with an ideal force source ($\underline{F}_0$) and an unknown mechanical impedance of the device $\underline{Z}_D$ is used.

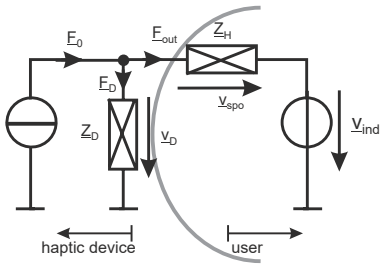


Fig. 4.19 Mechanical representation of the model of haptic interaction (fig. 4.3) with the example of a device combining an ideal force source and a complex impedance.

The nodal equation for the model (fig. 4.19) results in

$$\underline{F}_0 = \underline{F}_D + \underline{F}_{\mathrm{out}}\,. \tag{4.15}$$

By the usage of equation 4.14 the perception of an output force generated by a technical system gives the relation

$$\underline{K} = -\frac{\underline{F}_{\text{out}}}{s\,\underline{Z}_H}\,\underline{G}_{\text{FIP}}\,. \tag{4.16}$$

Equation 4.15 resolved according to $\underline{F}_{\text{out}}$ and put into equation 4.16 results in

$$\underline{K} = \frac{\underline{Z}_D\,\underline{v}_D - \underline{F}_0}{s\,\underline{Z}_{\text{H}}}\,\underline{G}_{\text{FIP}} \tag{4.17}$$

The movement velocity $\underline{v}_D$ of the device is the sum of $\underline{v}_{\text{ind}}$ and $\underline{v}_{\text{spo}}$, being identical to the velocity $\underline{v}_{\text{out}}$ in figure 4.3. The solution of the model in figure 4.19 according to equation 4.17 is adequate to quantify the haptic quality of any device resembling a force source. If the intention is to take the opposite approach - to calculate the design of a system based on the modelled interaction - equation 4.17 should be left dependent on $\underline{v}_{\text{ind}}$. This can be achieved most easily by a mesh equation of velocities

$$\underline{v}_D = \underline{v}_{\text{spo}} + \underline{v}_{\text{ind}}\,. \tag{4.18}$$

Replacing the device's velocity $\underline{v}_D$ with the impedance and integrating $\underline{v}_{\text{spo}} = \underline{x}_{\text{spo}}\,s$ results in

$$\frac{\underline{F}_D}{\underline{Z}_D} = \underline{x}_{\text{spo}}\,s + \underline{v}_{\text{ind}}\,. \tag{4.19}$$

Using equation 4.15 and with $\underline{F}_{\text{out}} = \underline{Z}_{\text{H}}\,\underline{x}_{\text{spo}}\,s$ equation 4.19 can be transferred to

$$\underline{F}_0 - \underline{v}_{\text{ind}}\,\underline{Z}_{\text{D}} = \underline{x}_{\text{spo}}\,s\,(\underline{Z}_{\text{D}} + \underline{Z}_{\text{H}}) \tag{4.20}$$

This equation - resolved according to $\underline{x}_{\text{spo}}$ and connected with the perception $\underline{K}$ by a variation of equation 4.16 - results in a complete description of an arbitrary haptic system with force source based on user-induced movements:

$$\underline{K} = -\underline{x}_{\text{spo}}\,\underline{G}_{\text{FIP}} = \frac{\underline{v}_{\text{ind}}\,\underline{Z}_{\text{D}} - \underline{F}_0}{s\,(\underline{Z}_{\text{D}} + \underline{Z}_{\text{H}})}\,\underline{G}_{\text{FIP}} \tag{4.21}$$

4.4.2 Tactile System

For a tactile system (fig. 4.20) a simplified model with one ideal velocity source $\underline{v}_0$ and an unknown mechanical device impedance $\underline{Z}_D$ is used.

The solution of this model (equ. 4.25) is applicable to quantify the haptic quality of any device which can be approximated by a velocity- or elongation source. The perception itself is dominated by a spontaneous elongation reaction based on the user's impedance. The corresponding velocities are

$$\underline{v}_{\text{spo}} = \underline{v}_0 - \underline{v}_{\text{ind}} - \underline{v}_{ZD}. \tag{4.22}$$

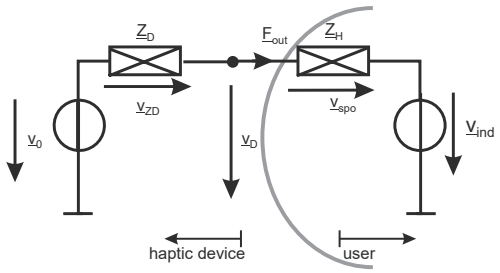


Fig. 4.20 Mechanical representation of the model of haptic interaction (fig. 4.3) with the example of a device combining an ideal velocity source and a complex impedance.

After integration and insertion into equation 4.9 the formula

$$\underline{K} = \underline{G}_{\mathrm{FIP}}\, x = \frac{\underline{v}_0 - \underline{v}_{\mathrm{ind}} - \underline{v}_{ZD}}{s}\, \underline{G}_{\mathrm{FIP}}. \tag{4.23}$$

becomes available. This one, somewhat more structured, and written in dependency of the output force $\underline{F}_0$ results in

$$\underline{K} = (\frac{\underline{v}_0}{s} - \frac{\underline{v}_{\mathrm{ind}}}{s} - \frac{\underline{F}_0}{\underline{Z}_D s})\, \underline{G}_{\mathrm{FIP}}\,, \tag{4.24}$$

and in a notation of elongations:

$$\underline{K} = (\underline{x}_0 - \underline{x}_{\mathrm{ind}} - \frac{\underline{F}_0}{\underline{Z}_D s})\, \underline{G}_{\mathrm{FIP}} \tag{4.25}$$

Equations 4.24 and 4.25 can frequently be simplified for common tactile systems, as

- with tactile systems, typically, frequency ranges are of relevance, in which the human user does not interact with the technical system making $\underline{v}_{\mathrm{ind}} = 0$,
- additionally the stiffness of the technical system is quite high.

The first assumption leads to a simplification of equation 4.24, i.e. to a much clearer relationship

$$\underline{K} = (\frac{\underline{v}_0}{s} - \frac{\underline{F}_0}{\underline{Z}_D s})\, \underline{G}_{\mathrm{FIP}}\,. \tag{4.26}$$

And with the second assumption equation 4.26 becomes as simple as

$$\underline{K} = \frac{\underline{v}_0}{s}\, \underline{G}_{\mathrm{FIP}} = \underline{x}_0\, \underline{G}_{\mathrm{FIP}}\,, \tag{4.27}$$

which reflects the origin of the transfer function $\underline{G}_{\mathrm{FIP}}$ as being a factor which is proportional to the perception of an elongation.

4.4.3 Examples in Time and Frequency Domain

Based on above models, figures 4.19 and 4.20 and the corresponding systematic descriptions with equations 4.17, 4.21 and 4.25, a number of questions concerning the design and analysis of haptic systems can be analyzed.

4.4.3.1 The Influence of Discretization of Force Signals on Haptic Perception

In many applications the force $\underline{F}_0$ is generated by a time-discrete system. Immediately the question arises whether time discretization is applicable and at which frequency the lower limits of the technical requirements are. Each discretisation of arbitrary dynamic signals generates steps in the rhythm of the output frequency. Such a step shows its maximum amplitude when the transmitted signal has its maximum gradient. Being of rectangular shape, it contains frequency components which diminish to the power of two starting at the discretization frequency. At a discretization of e.g. 100 Hz single steps are perceived very explicitly[6]. However, it is not immediately obvious at which frequency the ripples in the perceived force are small enough not to be perceived. Figure 4.21 shows the calculations based on the method presented before, for a base sinus of 2 Hz with amplitude of 2 N of a simple mechanical system. It is composed of a mass with 20 g and a friction coefficient of 0.1 Ns/m. The force is discretized with frequencies of 100 Hz and 10 kHz (a), resulting in the FFT (b) with a dominating base frequency at 2 Hz and peaks of the integer multiple of the discretization frequency. Within the perception domain $\underline{K}$ - filtered by the haptic dynamics - the force discretized with 100 Hz shows much smoother force steps and the discretization at 10 kHz does not show any artifacts at all (c). This is also reproduced in the FFT (d). This example shows distinctively the damping properties of the human haptic perception for high dynamic oscillations. By the application of the method described in the preceding chapters, a visualization and quantification of steps within time discrete systems is possible.

4.4.3.2 The Influence of Discretization-Width on the Haptic Perception

Another side effect of discretisation is the dead time resulting from the digital to analogue conversion. Dead times are of course nonlinear and therefore can only be approximated within the linear analyzing methods used here. Such an approximation is the description of the dead time T_t by a PT1 element with a time constant T_t according to

$$\underline{G}_T = \frac{1}{1 + T_T s}\,. \tag{4.28}$$

[6] One reason, why such a slow force output is not done at all

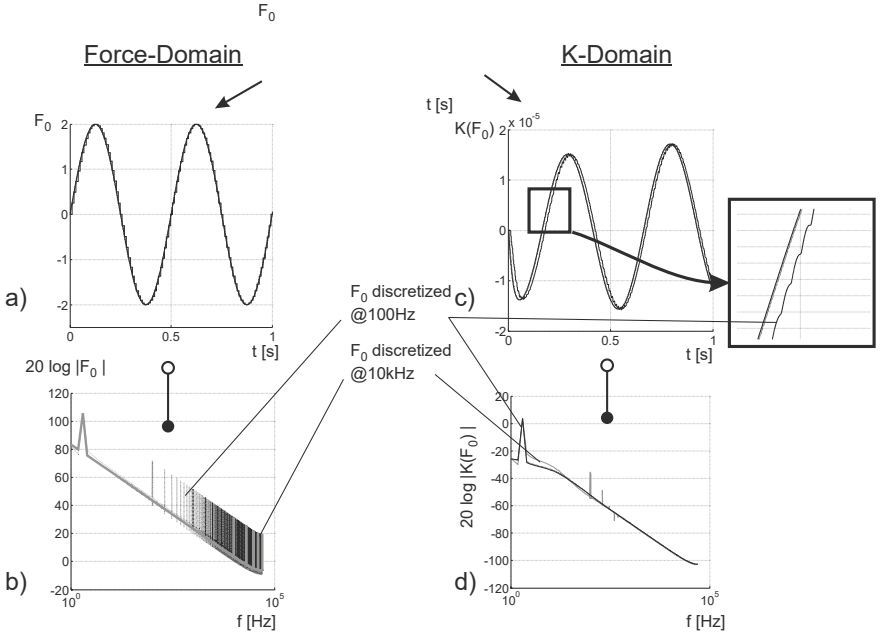


Fig. 4.21 Visualization in time- and frequency-domain of the haptic perception of a sinoid force, discretized with 100 Hz and 10 kHz.

In general, first-order lag elements within a signal conditioning chain are identical to a reduction of dynamics within the system. This effect can be observed practically when rendering hard contacts. A collision, e.g. in a simulation of two objects and the haptic feedback of the acting forces with a device, is perceived stiffer if the frequency range displayed is wider. System-engineers report that the upper limit of frequency range is not located at the typically assumed limit of tactile interaction of around 1 kHz. Instead they say that the rendering of calculated force with up to 10 kHz results in a definite increase in haptically perceived stiffness. This effect should possibly be confirmed by an analysis of the method shown here. The basis for the graphs given in figure 4.22 is the same technical setup as before. It is excited with a step of 2 N at 0.5 s. The discretization does not affect the amplitude of the system; the effect analyzed in the preceding example should not affect the quality of the perception. However, the PT1 element being the approximation of the dead-time element does not react as dynamically as it could. This is an effect which does not much affect the FFT of the physical force (b), but makes an obvious difference in the display of perception K in time domain (c) and frequency domain (d). The function discretized at 100 Hz has a very smooth but visible inclination. Compared to the input only slight differences can be observed with the functions discretized at 10 kHz and 1 kHz. The FFT (d) additionally shows that the frequency range at 10 kHz covers a 20 to 30 Hz more dynamic range. This is an effect which pays off

in tactile applications, especially as the differences lie exactly within the range of highest tactile sensitivity.

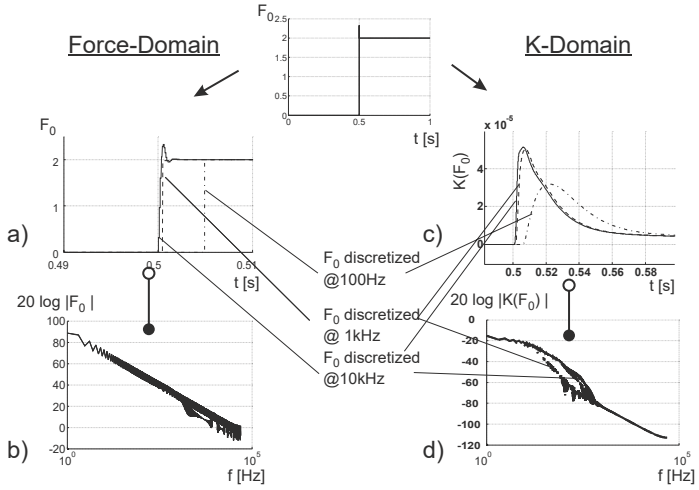


Fig. 4.22 Visualization in time- and frequency-domain of the haptic perception of a force step at time interval $t = 0.5\,\mathrm{s}$ in dependency from the discretization frequencies 100 Hz, 1 kHz, and 10 kHz .

4.4.3.3 Frequency Domain Analysis in the Design Phase

Another possible application of the above method lies in the frequency analysis during the design process of haptic systems. Figure 4.23shows a frequency response of the model of a kinaesthetic system according to figure 4.19. As an input a PT2 approximation of a user-induced velocity $\underline{v}_{\mathrm{ind}}$ with a border frequency of 10 Hz and an actuator output force $\underline{F}_0$ with a border frequency of 100 Hz was assumed. Additionally two cases were observed, one with the maximum output force $\underline{F}_0$ and another one with just $\underline{v}_{\mathrm{ind}}$ and the rest of the system being in an idle situation. The technical system includes a mass of 20 g and a perception threshold at -138 dB, a value which can be calculated by the approximation of the perception curve according to equation 4.10. The observed situation describes the excitation of a finger with a pestle in a large area (compared to the finger-tips area). Remarkably, even in idle state the perception curve K_{min} of this quite light and extremely dynamic system is below the perception threshold at frequencies below 2 Hz, only. The area between the curves K_{min} and K_{max} is a measure for the haptically available power of the system. It is a factor of the quality of a haptic device. Comparable visualisations of haptic devices allow analyzing the validity of a technical design in advance.

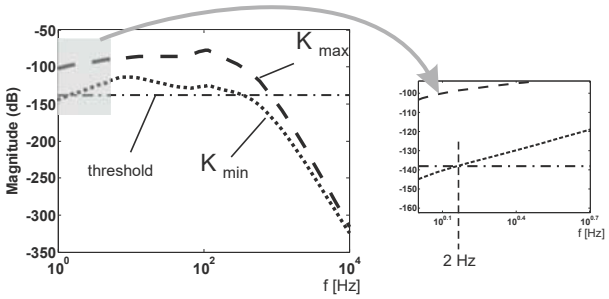


Fig. 4.23 Frequency plot of a kinaesthetic system according to equation 4.21 with maximum force-ouput and in idle mode.

4.4.3.4 Haptic Quality Estimations of Passive Haptic Devices

The method presented above may also be used to analyze the haptic quality of passive haptic devices like buttons and knobs; and even textures and surfaces. In a preliminary study: five visually identical, standard household light switches were chosen and modified in their haptic switching characteristics. By the exchange of two springs in the switch, the slope of the force-increase during down movement and the sharpness of the event of key press was modified (fig. 4.24).

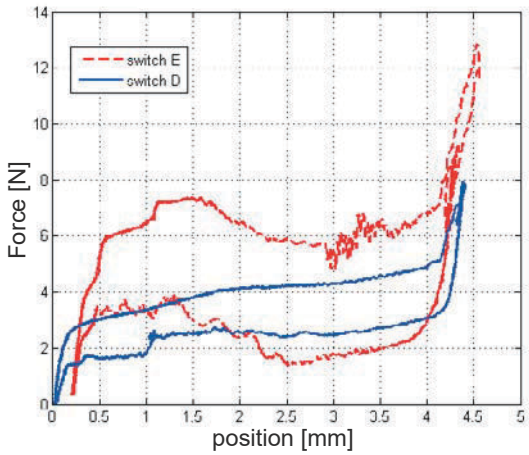


Fig. 4.24 Plot of force vs. position curve of two different switches.

The switches were presented to subjects asked to order them according to their haptic “explicitness” when pressed. This question was consciously chosen to be

unsharp in its meaning, as the subjects should only make themselves aware of and judge the haptic quality when pressing the switches. The subjects should not focus on anything specific - as this will not happen in e.g. a typical shopping situation, either. As a result a user rating of the five switches according to figure 4.25 came up. This user rating can be regarded as a reference to our further analysis.

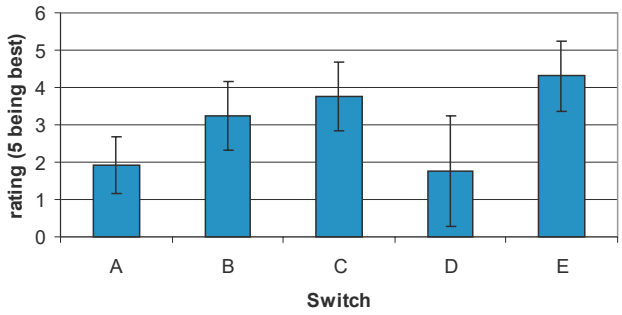


Fig. 4.25 Mean user rating for the haptic quantification of household light switches with standard deviation for five switches A-E.

For measurement purposes the switches were pressed at a frequency of 2 Hz with an electrodynamic actuator. Forces were measured with an ATI nano17 force/torque sensor. Displacements were measured with a Keyence LK-G32 laser triangulation system and sound pressure with a Brüel & Kjær condenser microphone. All switches were mounted on the measurement setup and pressed four times; excited with a trapezoidal force increase with 3 s period time. Force, displacement and sound pressure were sampled simultaneously with a sampling rate of 44 kHz. Except for a subsonic filter with a corner-frequency of 22.4 Hz in the microphone amplifier no filtering was used in recording the signals. Auditory noise was measured and removed from the recorded signal by subtracting the spectra respectively for future analysis of multimodal effects on perception. The measurements were analyzed in three different domains according to three characteristic values, each. Within the physical domain the absolute height of the tactile switching event (switch force), and the relative height between this point and the lowest point afterwards (switch difference) were identified (fig. 4.26a). Additionally, the click-ratio as the quotient of switch difference to switch-force was calculated. The click-ratio is a standardized way to characterize industrial switches according to DIN 41636. In a bandwidth-limited domain the original force-measurement data were filtered by a band-pass filter ranging from 10 Hz to 2 kHz, focusing on the range of tactile perception only. The maximum and minimum were identified within the resulting data (fig. 4.26b) giving three characteristic values: filtered maximum, filtered difference and filtered click-ratio.

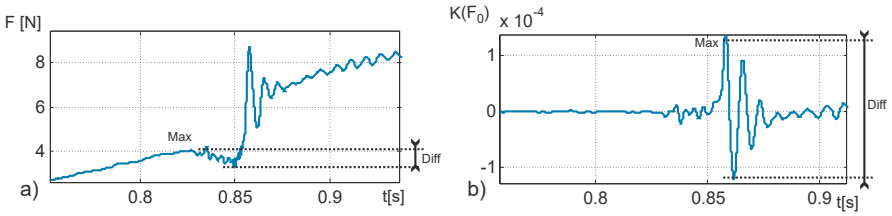


Fig. 4.26 Example for the identification of characteristic values in force-domain (a) and whithin the *K*-domain (b).

In the *K*-domain the bandwidth-limited data were additionally weighted according to the described method as stated in equation 4.14. The resulting data were analyzed identically to the bandwidth-limited domain. The identified values were normalized according to their maximum and set into relation to the normalized mean values in user ratings. All values were evaluated within a frequency range for the measurement data up to 2 kHz. Additionally, measurement data were artificially limited in bandwidth to 50 Hz, which makes them comparable to setups used in industry to measure force/displacement curves. Figure 4.27 gives the quadratic error sum of all five switches against user rating for all nine values.

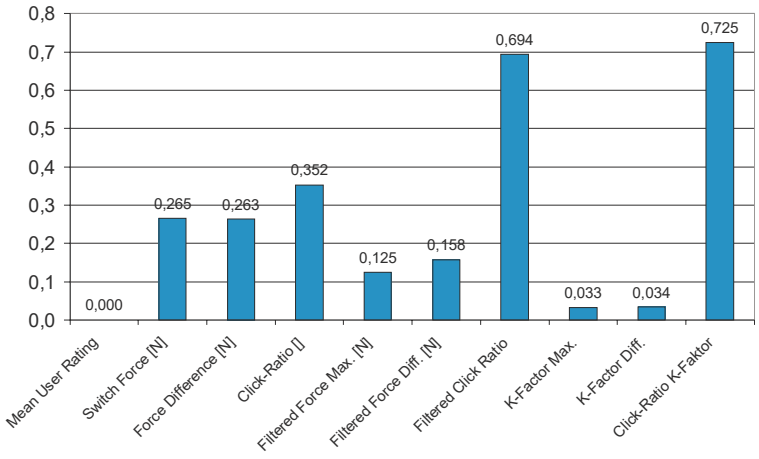


Fig. 4.27 Quadratic error between characteristic value and mean user rating for bandwidth limited measurement data with an upper border frequency of 2 kHz.

As a result any analysis including wide bandwidths gives a better rating than existing reference methods. Additionally, the *K*-domain analysis has the lowest

quadratic error compared to all other analyses.

This experiment proves the applicability of this method for the analysis of measurement data taken from passive haptic devices like switches.

4.5 Summarizing Remarks on the Application of the Method

According to the above method the frequency dependent parameter $\underline{K}$ is a measure for the haptic perception of mechanical oscillations. The quality and meaning of K depends largely on the quality of $\underline{G}_{\mathrm{FIP}}$ and $\underline{Z}_{\mathrm{H}}$, i.e. the performance of the models of the frequency dependency of perception and the mechanical load of the user. At the time of this book going to press both functions and their dependencies are subject to ongoing research activities. When using the presented method, we are recommended to make use of the most actual findings available. Special care should be taken of selecting the best fits of data of contact and grasp situations. In case such values are not available, the data presented in sections 4.2 and 4.3 are a good first fit for an approximation. Nevertheless, they should only be used with sensitivity to the context and the theoretical basis of this methodology.

"*This page left intentionally blank.*"

Chapter 5
Internal Structure of Haptic Systems

THORSTEN A. KERN; OLIVER MECKEL

In the preceding chapters the focus of the discussion was .haptic perception with regard to the human user. In the following chapter the technical realization of haptic systems will come to the fore. Consequently, the general view will change from a user centralized perspective to a device specific viewpoint. The understanding gained as well as the methods for the quantification of haptic perception will still be used for the analysis of the quality of technological solutions.

5.1 System Engineering

When starting the design of haptic devices, the engineer has to to deal with the general structures they can be composed of. Haptic devices of similar functionality can consist of very different modules. There are four big classes of possible system designs::

1. "open-loop admittance controlled systems"
2. "closed-loop admittance controlled systems"
3. "open-loop impedance controlled systems"
4. "closed-loop impedance controlled systems"

Impedance controlled systems are based on the transfer characteristics of a mechanical impedance $\underline{Z} = \frac{\underline{F}}{\underline{v}}$ and are typical of the structure of many kinaesthetic devices. They generate a force as output and measure a position as input. Admittance controlled systems instead, are based on the definition of a mechanical impedance $\underline{Y} = \frac{\underline{v}}{\underline{F}}$, describing a transfer characteristics with force-input and velocity-output. These systems generate a position change as haptic feedback and get a force reaction from the user as input source. In the situation of a closed-loop controlled system

T.A. Kern (ed.), *Engineering Haptic Devices*, DOI 10.1007/978-3-540-88248-0_5,
© Springer-Verlag Berlin Heidelberg 2009

this force is measured and used for the correction of the position. These analyses can be regarded as analog in the case of torque and angle replacing force and position for rotary systems. Nevertheless, for readability purposes, the following descriptions concentrate on translational movements and devices, only. In the subsequent chapters we will take a look at the details of all four variants.

5.1.1 Open-Loop Impedance Controlled

Open-loop impedance controlled systems are based on a quite simple structure (fig. 5.1). A force signal $\underline{S}_F$ is transferred via a driver $\underline{G}_{ED}$ into a force-proportional energy form $\underline{E}_F$. This energy is then altered into the output force $\underline{F}_0$ by an actuator $\underline{G}_{D1}$. This output force interferes with a disturbing force $\underline{F}_{noise}$. This noise is a result of movements generated by the user $\underline{x}_{out}$ and the mechanical properties of the kinematic design $\underline{G}_{D3}$. Typically, such disturbing forces are friction and inertia. The sum of both forces is the actual output force $\underline{F}_{out}$ of the impedance controller system. Usually, there is an optional part of the system, a sensor $\underline{G}_{D2}$, which measures the movements and the actual position of the haptic system.

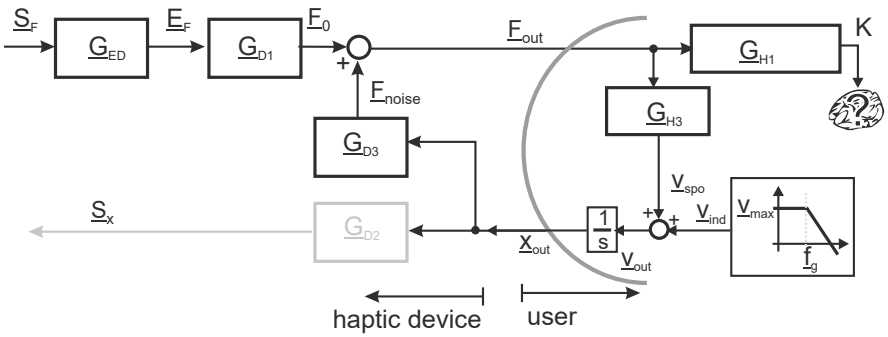


Fig. 5.1 Block-diagram of an open-loop impedance controlled haptic system.

Examples

Open-loop impedance controlled systems are the most frequently available devices on the market. As a result of their simple internal design, a few standard components can already be used to build a quite useful design with adequate haptic feedback, if at least some care is taken of the minimization of friction and masses. Among the cheapest designs available on the market today, the PHANTOM Omni (fig. 5.2a), connected via Fire-Wire to the control unit, is among the best known. It is frequently

used in research projects and for the creative manipulation of 3D-data during modeling and design. In the higher-price segment there are numerous products, e.g. the devices of the company Quanser. These devices are usually equipped with a real-time MatLab based control station, adding some flexibility to the modifications of the internal data processing by the end customer. The doubled Pantograph-kinematics of the "Haptic Wand" (fig. 5.2b) allows force feedback in up to five degrees of freedom with three translations and two rotations. Although all these devices may be open-loop impedance controlled, the software usually includes simple dynamic models of the mechanical structures. This allows some compensation of inertial and frictional effects of the kinematics based on the continuous measurement of position and velocities.

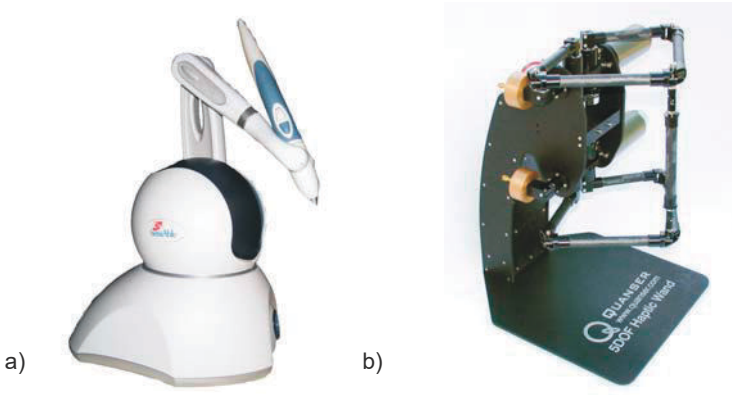


Fig. 5.2 Example of an open-loop impedance controlled system with a) serial-kinematic (PHANTOM Omni, SensAble) and b) parallel-kinematic (5 DOF Haptic Wand, Quanser) structure.

5.1.2 Closed-Loop Impedance Controlled

Closed-loop impedance controlled systems (fig. 5.3) differ from open-loop impedance controlled systems in such a manner that the output force $\underline{F}_{\text{out}}$ is measured by a force sensor $\underline{G}_{\text{FSense}}$ and is used as a control variable to generate a difference value $\Delta\underline{S}_{\text{F}}$ with the nominal value. An additional component typically is a controller $\underline{G}_{\text{CD}}$ in the control path, optimizing the dynamic properties of the feedback loop. The closed-loop makes it possible to compensate the force $\underline{F}_{\text{noise}}$ resulting from the mechanics of the systems. This has two considerable advantages: On the one hand, at idle state the system behaves in a much less frictional and more dynamic way compared to similar open-loop controlled systems. Additionally, as the closed-loop design allows some compensation of inertia and friction, the whole mechanical setup can be

designed stiffer. But it has to be noted that of course part of the maximum output power of actuators will then be used to compensate the frictional force, which makes these devices slightly less powerful than an open-loop design.

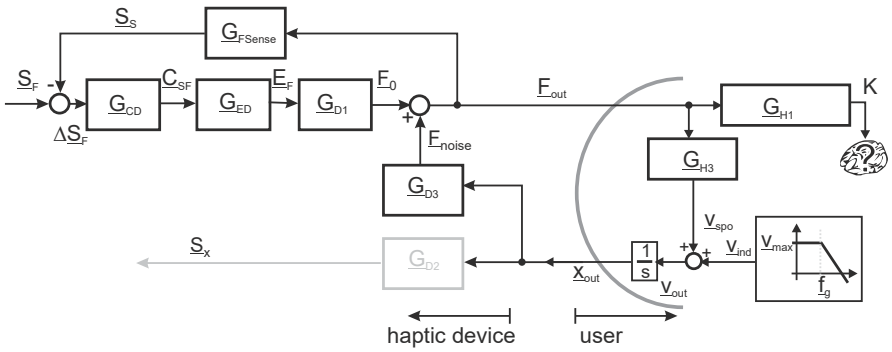


Fig. 5.3 Block-diagram of a closed-loop impedance controlled system with force-feedback.

Example

Closed-loop impedance controlled systems are usually used in research projects and as special purpose machines. The delta-series of ForceDimensions (fig. 5.4) is one example, as it is a commercial system with the option to buy an impedance controlled version. In this variant, force sensors are integrated into the handle, able to measure interaction forces in the directions of the kinematics' degrees of freedom. Closed-loop impedance controlled systems are technologically challenging. One the one hand they have to comply with a minimum of friction and inertia, on the other hand, with little friction, the closed loop tends to become instable, as an energy exchange between user and device may build up. This is why controllers, typically, monitor the passive behavior of the device. Additionally, the force sensor is a cost-intensive element, at present. In case of the delta-device, the challenge to minimize moving masses has been faced by a parallel-kinematics design.

5.1.3 Open-Loop Admittance Controlled

Open-loop admittance controlled systems (fig. 5.5) provide a positional output. Proportionally to the input value $\underline{S}_x$, a control chain with energy converter $\underline{G}_{ED}$ and kinematics $\underline{G}_{D1}$ provides a displacement $\underline{x}_0$. This displacement interferes with a disturbance variable $\underline{x}_{noise}$ which is dependent on the mechanical properties of the kinematics $\underline{G}_{D3}$ and a direct reaction to the user's input $\underline{F}_{out}$. In practice an open-



Fig. 5.4 Example of a parallel-kinematic closed-loop impedance controlled system (delta3, ForceDimension).

loop admittance controlled system typically shows a design which allows to neglect the influence of the disturbance variable. Another optional element of open-loop admittance controlled systems is the measurement of the output force with a force sensor $\underline{F}_{\mathrm{Sense}}$ without closing the control loop.

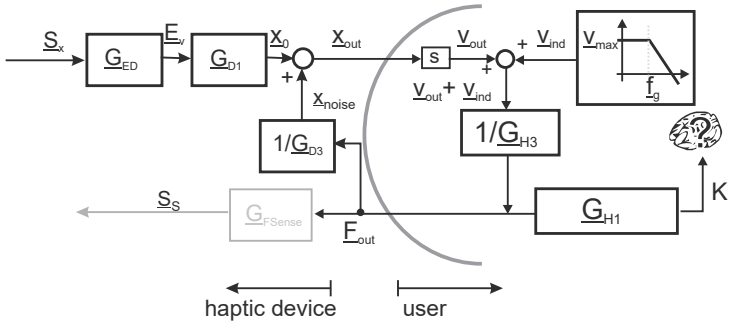


Fig. 5.5 Block-diagram of an open-loop admittance controlled haptic system.

Example

Open-loop admittance controlled systems are used especially with haptic devices in the area of tactile displays. Many tactile displays are based on pin arrays, meaning that they are generating spatially distributed information by lifting and lowering pins out of a matrix. These systems' origins are Braille devices (fig. 5.6) coding letters

in a tactile, readable, embossed printing. For actuation of tactile pin-based displays a variety of actuators are used. There are electrodynamic, electromagnetic, thermal, pneumatic, hydraulic and piezoelectric actuators and even ultrasonic actuators with transfer media.

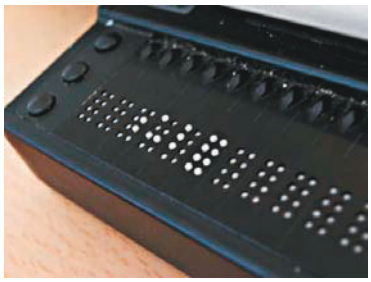


Fig. 5.6 Example of an open-loop admittance controlled system being a Braille row (Wikipedia).

5.1.4 Closed-Loop Admittance Controlled Devices

Closed-loop admittance controlled devices (fig. 5.7) provide a positional output and a force input to the controlling element identical to impedance controlled devices. The mandatory measurement of the output force $\underline{F}_{\text{out}}$ is used as control variable $\underline{S}_{\text{S}}$ for calculating the difference $\Delta\underline{S}_{\text{F}}$ with the commanding value $\underline{S}_{\text{F}}$. This difference is then fed through the controller $\underline{G}_{\text{CD}}$ into the control circuit. As a result, the displacement $\underline{x}_{\text{out}}$ is adjusted until an aspired force $\underline{F}_{\text{out}}$ is reached.

A variant of a closed-loop admittance controlled device is shown in figure 5.8. Closed-loop admittance controlled devices show considerable advantages for many applications requiring large stiffnesses. However, the force sensors $\underline{G}_{\text{FSense}}$ are quite complex and consequently expensive components, especially when there are numerous degrees of freedom to be controlled. As a variant, the system according to figure 5.8 does not use a sensor but just a force-proportional measure, e.g. a current, as control variable. When using e.g. a current with electrodynamic actuators, we can identify even the reaction of the user generating an induction as an additional velocity dependent value.

Examples

At present, closed-loop admittance controlled systems are the preferred approach to provide high stiffnesses with little loss in dynamic properties. The idea to hap-

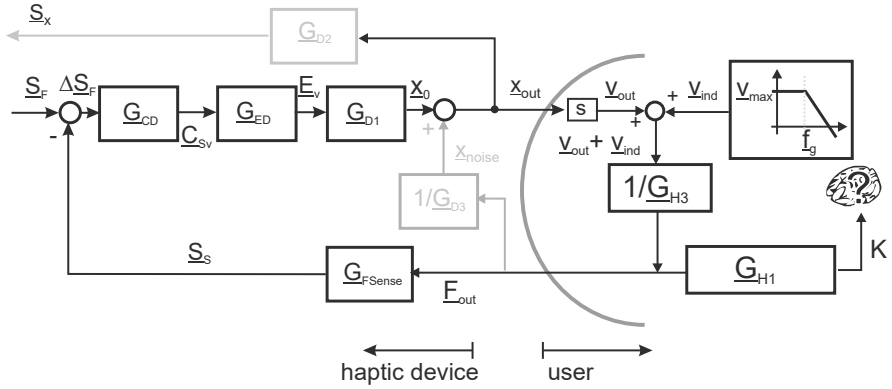


Fig. 5.7 Block-diagram of a closed-loop admittance controlled haptic system with force-feedback loop for control.

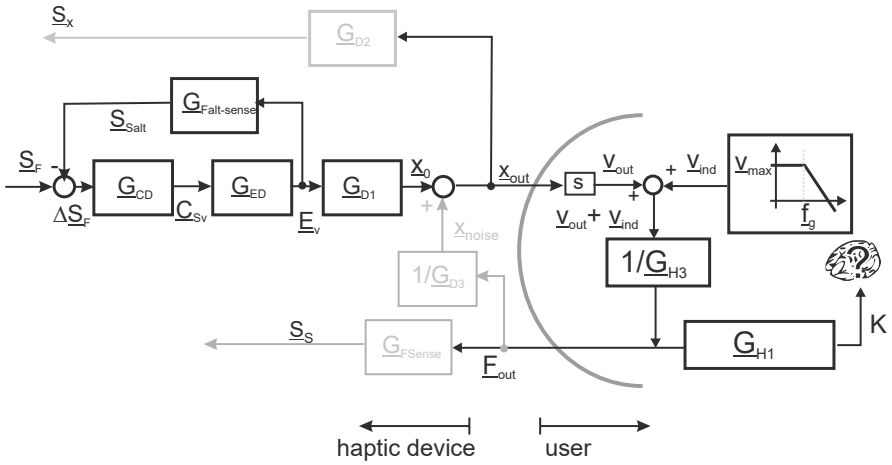


Fig. 5.8 Block-diagram of a closed-loop admittance controlled haptic system with a feedback loop measuring an internal force-proportional value.

tically hide the actual mechanical impedance from the user by closing the control loop makes it possible to build serial kinematics with a large workspace. The FCS HapticMaster (fig. 5.9a) is such a one meter high system with three degrees of freedom and a force of up to 100 N. It includes a force sensor at its handle. The axes are controlled by self-locking actuators. The device's dynamics is impressive, despite its size. However, a damping has to be included in the controller for security reasons resulting in a limitation of bandwidth depending on the actual application. Realisations of the variant of closed-loop admittance controlled devices are the Virtuose-systems from Haption (fig. 5.9b). In these devices the current is measured at elec-

trodynamic electronic commutated actuators and fed back as a control value. The devices show impressive qualities for the display of hard contacts, but have limited capabilities in the simulation of soft interactions, e.g. with tissues. Therefore, the application area of Haption-systems is mainly the area of professional simulation of assembly procedures for manufacture-preparation.



a)



b)

Fig. 5.9 Examples of closed-loop admittance controlled systems in variants with a) direct force measurement (HapticMaster, Moog) and b) measurement of the actual current (Virtuose 6D35-45, Haption).

5.1.5 Qualitative Comparison of the Internal Structures of Haptic Systems

As the haptic human-machine interaction is based on an impedance coupling, it is always the combination of action and reaction, be it via force or position, which has to be analyzed. In fact, without any knowledge about the internal structure of a device, it is impossible to find out whether the system is open-loop impedance controlled, closed-loop impedance controlled or closed-loop admittance controlled. With experience of the technological borders of the most important parameters like dynamics and maximum force, an engineer can make a well-founded assumption about the internal structure by simply using the device. But concerning the abstract interface of in- and output values, all the devices of the above three classes are absolutely identical to the user as well as to the controlling instance. Despite this fact the technical realizations of haptic systems differ widely in their concrete technical design of course. The parameters influencing this design have to be balanced against each other. Such parameters are:

- Number of components
- Maximum impedance to be achieved at slow motion

- Minimum impedance to be achieved at fast motion
- Force-resolution
- Impedance of mechanical components (e.g. inertia of kinematics)

These parameters and their mapping onto the technical designs are given qualitatively. In figure 5.10 the impedance generated by a device in absolute values and the impedance range covered may be one criterion for the performance of a device. Analyzing the systems according to this criterion shows that open-loop admittance controlled systems may have high impedance, which shows smaller variability in tighter borders. Closed-loop admittance controlled systems extend these borders by their ability to modulate the impedance due to the feedback loop. Depending on the design, closed-loop admittance controlled systems vary in the width of this modulation. In the lower area of possible, realizable impedances the open-loop impedance controlled systems follow. They stand out more by simplicity in their design than by large impedance ranges covered. In comparison to the closed-loop admittance controlled systems they gain some impedance width at the lower border of impedances. In order to be able to equally cover lower as well as higher impedances, the choice should be made of closed-loop impedance controlled systems.

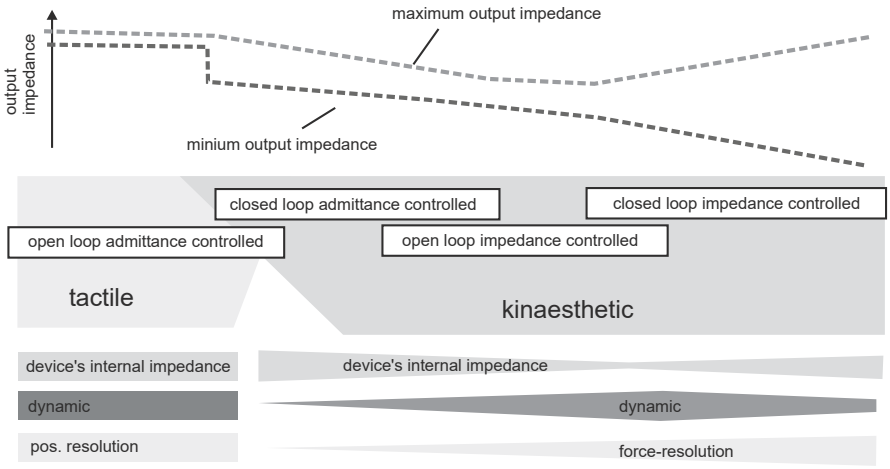


Fig. 5.10 Qualitative comparison of the application areas for different device-structures.

Normally, pure open-loop admittance controlled systems are suitable for tactile devices only, as, with tactile devices, usually there is no active feedback by the user to be measured. The haptic interaction is limited to tensions being coupled to the skin of the user's hands. Such devices show high internal impedance ($\underline{Z}_D$). The dynamics and the resolution concerning the displacement are very high.

Kinaesthetic devices can be built with systems allowing a modulation of the displayed impedances. The closed-loop admittance controlled systems excel due to the possibility to use mechanical components with high impedances. The dynamics of these systems are accordingly low (<100 Hz) and the force-resolution is, due to the typical frictions, not trivial when realized. Open-loop impedance controlled systems show a wider dynamic range due to the missing feedback loop with, at the same time, limited dynamic range. Only closed-loop impedance systems allow covering a wide impedance range from lowest to very high impedances., whereby with increasing requirements of force resolution the dynamics of the maximum velocities achieved by the control loop are limited and limitations of the measurement technology become noticeable.

The decision on the design of a haptic system has significant influence on the application range and vice versa. On the one hand, it is necessary to identify the requirements to make such a decision. For this purpose, it is necessary to ask the right questions and to have an insight into possible technical realizations of single elements of the above structures. This is the general topic of the second part of this book. On the other hand, it is necessary to formulate an abstract system description of the device. An introduction of how to achieve this is given in the following section.

5.2 Abstraction of the System Components

In the preceding section the general structures available for haptic systems have been introduced. In this section an abstraction of the individual system structure will be given as a basis for the design and analysis of the system's transfer characteristics. The content of this section will be further elaborated and detailed in chapter 7.

5.2.1 Basic System Analysis

All the approaches presented for the design of haptic systems, may they be impedance- or admittance controlled, can be abstracted and broken down into their major system components. For the overall system with its borders the complete mechanism and electronics of a haptic device are regarded. The basic idea behind the haptic system is to give a human operator, the user, an impression of the load and reactions of an apparatus or process. Yet in an almost trivial start of analysis, all necessary parts relevant for the structural analysis should be mentioned:

- a haptic device, - whether impedance or admittance based - providing a haptic feedback to the user on a value within a process,
- the process the user is going to influence and to change by specific actions,

- the user who must be regarded as being part of the overall system and who is coupled to the process by his reactions to the haptic feedback.
- a telemanipulator which supports the user in his influence on this process. This element may be totally part of a simulation and in such a case existing only in a time-discrete simulation engine.
- interface modules with signal conditioning and internal control circuits, interfacing the haptic device by simulated or measured process values according to defined conditions.
- additional control structures on a hidden level which will have an important role to play in guaranteeing the desired functioning of the system.

Each component itself is coupled to the overall system via interfaces, whereas these interfaces generate and receive system variables. Figure 5.11 extends the focus on a complete telemanipulation system.

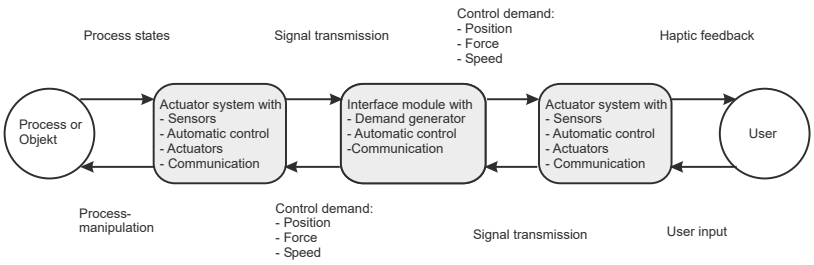


Fig. 5.11 Abstract representation of a telemanipulation-system.

A deeper analysis of the components in a telemanipulator is done on the basis of a known transfer characteristics of each system component. In the following section an approach will be given showing how to understand the different system components and describe them systematically.

5.2.2 Analysis of System Components

Within the total telemanipulation system there are numerous coupled system elements, differing significantly in their transfer characteristics. Each of them offers different opportunities and challenges as a result of the physical phenomena they are based on. In comparison with all other components, actuators of the haptic device or the manipulator offer the best chance to be described by mathematical formulations, e.g. differential equations. A general approach is given in the block structure of figure 5.12.

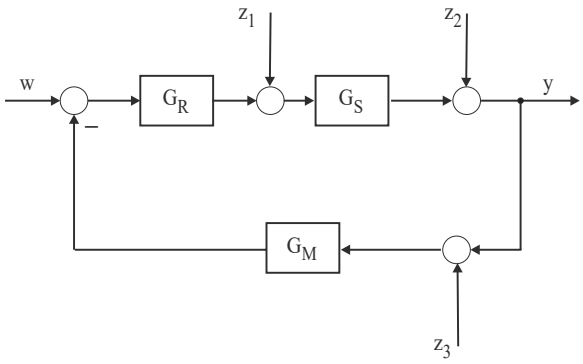


Fig. 5.12 Block-diagram of a closed-loop controlled actuator system.

For this example the actuator is placed in a closed control circuit. The control transfer function

$$\underline{G_F} = \frac{\underline{G_R}\underline{G_S}}{1 + \underline{G_R}\underline{G_S}\underline{G_M}} \tag{5.1}$$

totally describes the behavior of a controlled actuator. The transfer function of the controller is given by $\underline{G_R}$, the transfer function of the actuator by $\underline{G_S}$ and the transfer function of the sensor, e.g. a force sensor within admittance controlled systems, by $\underline{G_M}$. This sensor is placed in the feedback loop of the control circuit. Beside the in- and output values - the desired value or reference variable w and the controlled variable y - there are numerous disturbing variables z_i affecting this small subsystem. Dependent on the subsystem they have different origins and influences. In this example the disturbing variables are additive. Frequently nonlinear influences and dependencies on other values not directly part of the control circuit can be described by such disturbance variables. When considering many details of disturbing variables, we can see that a highly complex description evolves from even simple models, which sometimes hinders further analysis instead of helping to understand the system elements. It is necessary to choose a good balance between a detailed and an efficient description of the modeled system.

The descriptions of different principles of the generation of haptic feedback give an insight into the wideness of the solution space. Nevertheless, taking individual solutions and reducing them to their abstract description in a control engineering sense makes them surprisingly similar again:

- The general behavior of control systems, especially of electromechanic actuators, can frequently be reduced to a proportional value. For example an electric actuator is a transformator between an electric current I as input and a mechanical torque M as output. With a constant current as input, a constant torque as

output will be generated. In case of a second closed-loop circuit being placed around this actuator, e.g. with the task to reach a defined position, the whole system will show an integral behavior. Such a simple example already includes many elements of a system description. Beside the transfer characteristics of the pure actuator, elements of the power train have already to be considered.

- All systems show characteristic dynamic properties. This is a direct result of the inevitably existing energy storages (in a mechanical sense these are inertia and elasticities, in an electrical sense these are capacities and inductance). These energy storages have a direct impact on the system's response characteristics, which can be approximated roughly by the usage of delay elements of *n*'th order.
- Easily distinguishable from the delay elements the system components show a dead time. The impact of dead time elements on closed circuits is serious and can frequently be observed with haptic devices as a result of discrete signal processing (section 4.4.3). Transmission time of signals is vital to specific tasks of haptic research, e.g. with medical teleoperation and haptic feedback or telemanipulation over large distances in space or combat. More classic and trivial examples of obvious dead time elements like conveyor belts are very rare in haptic applications.

The term *actuating system* already implies the assumption that the actuating unit is controlled via an electronic circuit and that it is part of a superordinate hierarchy level. The typical inner disturbance variables of such an electromechanic system (varying frictional states, temperature dependency on electrical resistances, variations of source voltages, etc.) are regarded as being compensated within the system itself. In addition, new disturbance variables appear, resulting from the mechanical task itself (process loads, counter forces, positional changes, etc.). The disturbance variables can be modeled differently:

- An additive formulation is a comparably simple way to model the disturbance variable.
- If the process variable is known to be nonlinear and can be modeled by a nonlinear formulation the whole system is going to become nonlinear. Analysis and descriptions of nonlinear systems are discussed in chapter 7.

The actuators in a completely mechanical telemanipulator, are, on the one hand, part of the end-effector, on the other hand part of the haptic system. As a consequence, the interface structures of both applications are different. Figure 5.13 visualizes such an interface for the end-effector part with signal conditioning and control unit.

As its input, the actuating unit at this end uses a position as commanded value - typically an open-loop admittance controlled device with force measurement. It may also be structured as closed-loop impedance or admittance controlled element with force input. However, such a structure is mainly used with lightweight soft robotic manipulators and still is quite seldom. The output values are given by measured actual values and/or control deviations and/or inner system variables. As already

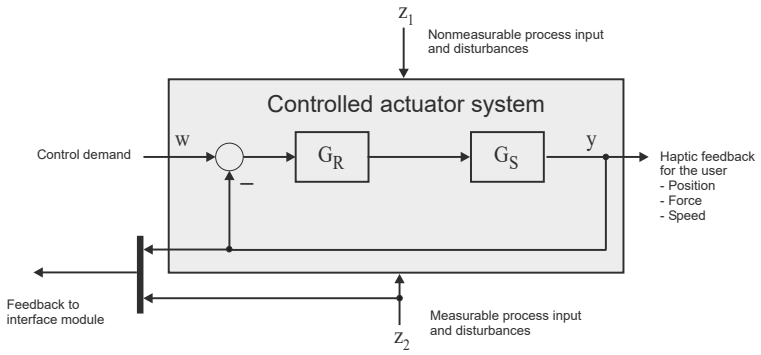


Fig. 5.13 Actuator-system in a telemanipulator.

mentioned the actuator system experiences additional disturbance and input values *Z* due to being integrated into a complex system.

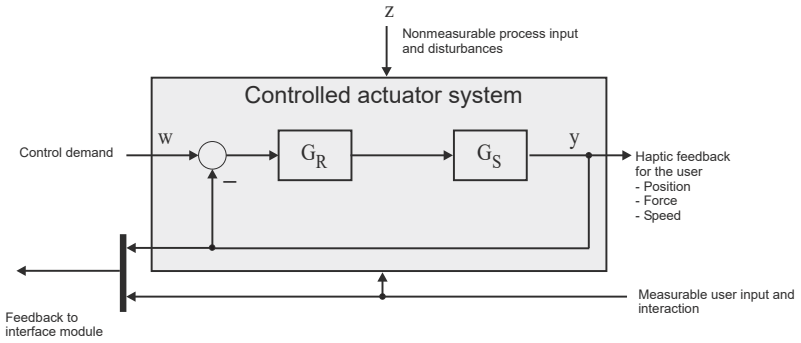


Fig. 5.14 Actuator-system in a haptic device.

Figure 5.14 shows the interface of an actuator system in a haptic device. The interface of the central control module for data processing is similar to the one between telemanipulator and control unit. Once again commanded values are transmitted, in the case of a haptic device they are forces, typically. Additionally, measured values, e.g. velocities or positions, are fed back to the controlling unit.

The central controlling unit between these actuating units is a node connecting the haptic device and the end-effector or its time-discrete simulation to a telema-

nipulation system. It is the core of the complete setup with a number of important functionalities:

- Commanding-value generation for the end-effector - an interpretation of the measured values acquired from the user has to take place. An analysis of the measurement may happen to filter unwanted spectral components which may excite the system in resonance areas. Other processing of the data may happen with the aim to keep the control loop stable.
- Commanding-value generation for the haptic device - an interpretation of the measured (or simulated) values acquired from the end-effector has to take place. During this value generation mathematical and logical operations may be used to modify this parameter. Any actual or historic values available within the system may be helpful to achieve the general aim: minimizing control deviation without its becoming instable.
- If there is a need to have access to as many actual values as possible, it may be helpful to integrate parts of the control circuits, from near the actuator, into the central controlling unit.

Given these functions, this central control and signal-conditioning module is an important part of the total system. Its transfer characteristics have to be subject to a design process identical to all other components, from actuators to sensors and electrical and software interfaces. However, the solution space is almost boundless making a systematic approach mandatory!

The peripheral surroundings of the telemanipulation system, e.g. the manipulated object or the subject, is reproduced by the process and the type of interaction. The preceding sections 4.2 and 4.3 provide a model of the user as a mechanical load with respect to the grasping situation and the grasping force. However, this model focuses on the higher frequency range of passive haptic interaction. For stability issues the lower frequency range in the area of below 10 Hz becomes relevant. The description of the user in this range of active, haptic interaction can only be done to a very limited extent. We recommend dealing with the user as a disturbance variable with defined amplitude and dynamic range, which can be guessed by an analysis of the task and user's typical performance. Nevertheless such an approximation has to be handled with care when being used as the basis for a design process and for the dimensioning of transfer characteristics of the complete system. Therefore more advanced systems monitor the device's behavior and adjust control values depending on its state (subsection 7.3.3). Describing the process and/or the manipulated object - in direct comparison - is in many cases much easier to be achieved. There are of course situations of different complexity depending on the universality intended to be achieved by the telemanipulator. However, most object interaction processes will be nonlinear at least, as grasp and collision and release actions include instant changes of load situations.

Simple examples of well known situations are e.g. the movement of masses within free space, which can be easily modeled with a differential equation. However, the formulation of resistive force of the propulsion of bore rods for gas and oil

exploration is almost unpredictable and subject to sudden changes depending on the geological layers passed. A help for modeling may be force measurement principles as discussed in chapter 10, which may be used not only for measurement during telemanipulation, but may also be part of measurement setups to optimize mathematical models of mechanical objects. The design of peripheral influence values is intended to support the major system components in the telemanipulation system. As a general statement, the design of the transfer functions of any component, may it be the central control and signal conditioning unit as being described in detail in chapter 12 or the description of disturbance variables will not be able to cover reality in its full complexity. The challenge is to create a model covering major influence parameters of the selected components. This demands an insight into the requirements of the task, the interaction being performed, and the limitations and possibilities inherent in the technical components. How to gain the knowledge necessary to achieve this will be the subject of the second part of this book.

Part II
Designing Haptic Systems

"*This page left intentionally blank.*"

In the first part of the book the basics of a general understanding of the design of haptic systems are given. There, the emphasis is on the discussion of the biological background and on a closer look at resulting technical questions. In the second part more definite technological subjects are examined and more concrete help for frequent challenges of the design process is offered. The chapters in this part are ordered according to the classic task list to be realized during any technical design process. They start with more general questions concerning the overall system and proceed to specialized questions concerning specific subcomponents. The chapters are intentionally ordered in such a way that the ones dealing with questions whose solution spectrum is limited to a great extent are discussed earlier than those which provide more flexible solutions applicable to many situations.

- Chapter 6: The acquisition of **requirements** for the technical design process is discussed. The design of haptic systems covers a plurality of technological questions. Especially the challenges concerning the high dynamics to be achieved make a systematic approach mandatory for identifying the requirements.
- Chapter 7.1: After the basic requirements have been identified, a superordinate view of the **structure of the system** to be built is necessary for which a methodology is given in this chapter. The resulting analysis does not only aim at the decision on the control structure of the device, but also defines the technological subproblems to be addressed during the following design process.
- Chapter 8: Especially kinaesthetic, but also multidimensional tactile systems have to combine multiple degrees of freedom to fulfill the requirements for certain tasks. This leads to the systematic design of the **kinematics** of the device. This chapter provides the necessary knowledge on kinematic design and covers specific and sometimes surprising problems of mechanical transfer functions for haptic devices in parallel kinematics.
- Chapter 9: In this most comprehensive chapter of the second part, typical **actuator principles** with respect to their application in haptic devices are discussed. The sections cover all popular actuation principles in an overview, as well as the details of selected principles for a design from scratch.
- Chapter 10: Especially closed-loop controlled haptic devices need **force sensors**. Furthermore, telemanipulation systems - besides simulators - are the second most relevant group of high fidelity haptic devices. A haptic telemanipulator requires force sensors at its end-effector, at least. This chapter covers the selection process and the design of force sensors according to the physical principles lying in their specifications within a haptically relevant range.
- Chapter 11: For a complete haptic interaction each system requires a **position measurement**. Technological solutions for this subordinate technological challenge are discussed in this chapter, whereby different positioning- and other movement-sensors are presented.
- Chapter 12: Typically, haptic devices are interfaced with time discrete systems with digital signal processors, may they be controllers for universal interfaces or complete simulation systems. This chapter deals with the **interfaces** between

these computing units and gives insight into the performance of different realizations.

- Chapter 13: The most frequent application of complex haptic systems is the perception of virtual realities in **simulations**. For this purpose, it is necessary to calculate the simulated physical properties by a software system running on a programmable device. This chapter deals with a selection of algorithms used for such calculations and gives an overview of the technological challenges. It also provides a terminological basis for mechanics and hardware developers in order to make them understand software engineers.
- Chapter 14: **Final remarks** on all previous chapters are made and the most important recommendations for the design of haptic systems are summarized.

Chapter 6
Identification of Requirements

THORSTEN A. KERN

6.1 The Right Questions to Ask

At the beginning of a technical design process the requirements for the product which, usually, are not clear and unambiguous, have to be identified. Frequently, customers formulate wishes and demands respectively solutions instead of requirements. A typical example is a task of the kind: "to develop a product just like product **P**, but better/cheaper/nicer". If an engineer accepts such a kind of order without getting to the bottom of the original motivation the project will be doomed to failure. Normally, the original wish of the customer concerning the product has to fulfill two classes of requirements: The product shall have

- a certain **function**
- in a distinct **technical** and **market oriented** framework

The formulations of the market oriented requirements are manifold yet not in the focus of the following analysis (for details of a general systematic product development see [189]). They may be motivated by an existing product **P** to compete with, but usually they are much more comprehensive and cover questions of budget, timeframe of development, personal resources and qualifications and customers to address.

With regard to the technical framework, the customer typically gives just unspecific details. A statement like "a device shall provide a force on a glove" is not a definition of a requirement but already a solution on the basis of existing knowledge on the part of the customer. The complexity of a real technological solution spans from a single actuator to provide e.g. a vibration to complex kinematics addressing single fingers. If you question the customer's original statement, it may even be that his intention is, e.g. to simulate the force impression when switching the gears of a

T.A. Kern (ed.), *Engineering Haptic Devices*, DOI 10.1007/978-3-540-88248-0_6,
© Springer-Verlag Berlin Heidelberg 2009

clutch in a passenger car. The knowledge about the actual application and following that knowledge about the interaction itself allows the developer a much broader approach, leading to a more optimized technical solution. The following sections identify classes of possible technical realizations and collect the necessary questions in clusters of possible applications. The list does not make a claim to be complete, but is the essence of requirement specifications of dozens of developments achieved during the last few years.

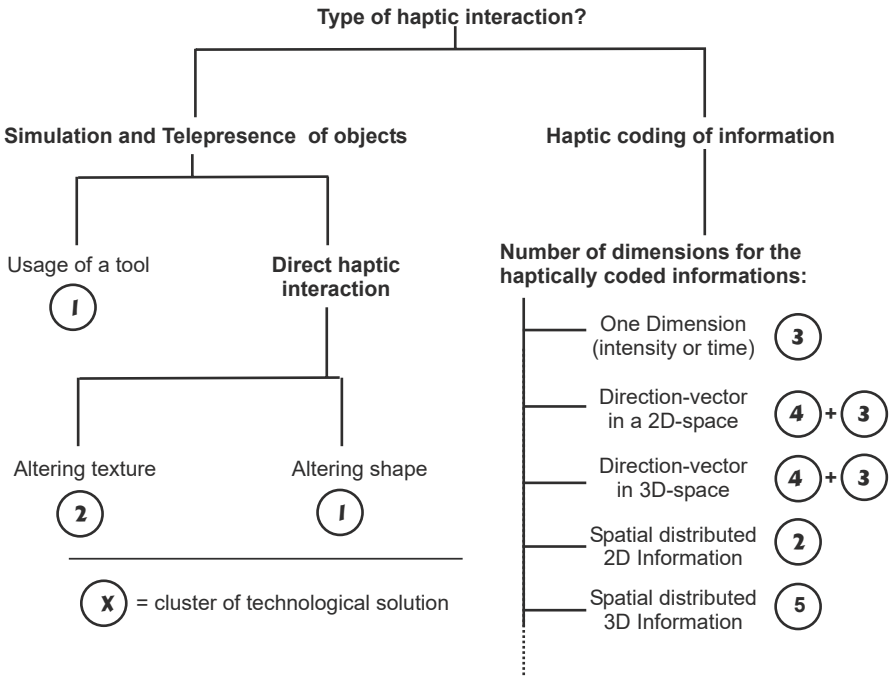


Fig. 6.1 Structure for identifying relevant clusters of questions by analyzing the intended haptic interaction.

In figure 6.1 a decision tree for the identification of clusters of questions has been sketched. It is recommended to follow the tree from top to bottom in order to identify the correct application and the corresponding cluster of questions. Subsection 6.1.1 describes the usage of the decision tree, whereas in the following sections the clusters of questions are highlighted in detail.

6.1.1 Interaction as a Classification Criterion

The core of the requirements' identification is the definition of the type of haptic interaction.

Type of Haptic Interaction

The first question asked should always refer to the type of interaction with the technical system. Is it a simulation of realistic surroundings, the interaction with physically available objects in terms of telepresence; or is the focus of the interaction on the pure communication of abstract information? In the former cases the variants are less versatile than in the latter.

Simulation and Telepresence of Objects

Does the interaction aim at touching virtual or via telepresence available objects? If this is the case, does the interaction take place directly via fingers, hands or skin, or is a mediator, e.g. a tool the interacting object? Does the user hold a specific tool - a pen, a screw driver, a surgical instrument, a joystick of a plane, in his hands and control one or more other objects with it, or does the user touch a plurality of objects during the interaction with his or her hands? In the case of a tool-interaction the chosen solution can be found in cluster ① "kinaesthetics", in the case of a direct interaction another detail has to be considered

Direct Haptic Interaction

By touching physical objects, the user can notice the differences in all physical attributes of the volume like mass, elasticity, plasticity and inner friction, and their texture (chapter 2). In the case of interacting with shapes, the questions of cluster ① "kinaesthetics" remains relevant, in the case of interacting with textures the questions of cluster ② "surface-tactile" have to be considered. This is not necessarily an alternative decision, however, with the same object interaction, both aspects can be required at the same time or one after the other.

Haptic Coding of Information

In the case of abstract, not physical, object- oriented information communication via the haptic sense the question of the dimension of information becomes relevant:

- Does the interaction include a single event which occurs from time to time (e.g. the call of a mobile phone) or is some permanently active information (e.g. a distance to a destination) haptically communicated? These questions are one-dimensional[1] and covered by cluster ③ "vibro-tactile".
- Is the interaction dominated by directional information coding an orientation in a surface (directional movement) or in a space? In this case the questions covered by cluster ④ "vibro-directional" are relevant. In such applications frequently time respectively distal information is included, also making the questions in cluster ③ become relevant.
- Does the interaction aim at the communication of data distributed within a two-dimensional information layer, like geological maps, road-maps or texts on a page? In these cases the questions of cluster ② "surface-tactile" have to be answered.
- In case there is volumetric information - the electrical field of an atomic bonding or medical data sets - to be haptically transmitted, the questions of cluster ⑤ "omnidimensional" are to be be considered.

In the following section the questions in the clusters are further discussed and some examples are given for the range of possible solutions to the questions aimed at.

6.1.2 Cluster "Kinaesthetic"

Cluster ① has to be chosen either when an interaction between fingers and shapes happens directly or when the interaction takes place between tool and object. Both cases are technical problems of multidimensional complexity[2]. Each possible dimension movement corresponds to one degree of freedom of the later technical system. Therefore the questions to be asked are quite straightforward and mainly deal with the requirements for these degrees of freedom of tools and users:

- Which degrees of freedom do the tool /movement show? → rotatory, translatory, combinations [3]
- How large is the volume covered by the movements? → Maximum and minimum values of angles and translations
- How dynamic is the movement? → Identification of maximum and minimum velocities and accelerations. Usually, this question cannot be answered immediately. A close observation of the intended interaction will help and - as far

[1] as the information includes only one parameter

[2] A tool interaction can be a one-dimensional task, but such an assignment concerning the technical complexity can be regarded as an exception

[3] In the case of a finger movement it has to be noted that not necessarily all movement directions have to be equipped with haptic feedback to provide an adequate interaction capability. Frequently it is even sufficient to provide the grasp-movement with haptic feedback, solely.

as possible - measurements of movement velocities of instruments and fingers should be made e.g. with the aid of videos,.
- Which forces / torques happen at each degree of freedom?[4]? → Definition of maximum and minimum forces and torques.
- What is the maximum dynamics of the forces and torques? → Bandwidth of forces and torques in frequency range, alternatively maximum slope in time-domain ("from 0 to F_{max} in 0.1 s"). Usually, this question cannot be answered directly. Frequently, measurements are difficult and time-consuming, as an influence of the measurement on the interaction has to be eliminated. Therefore it is recommended to do an analysis of the interaction itself and the objects the interaction happens with. If it is soft - something typical of surgical simulation, simple viscoelastic models can be used for interpolating the dynamics. The most critical questions with respect to dynamics often address the initial contact with an object, the collision. In this case especially the stiffness of the object and the velocity of the contact describe the necessary dynamics. But is has to be stated that the resulting high demands are not seldom in conflict with the technical possibilities (section 3.3).

6.1.3 Cluster "Surface-Tactile"

Haptic texture represents the microstructure of a surface. The lateral movement between this microstructure and the finger tip results in shear forces generating the typical haptic impression of different materials. Haptic-bumps on the keyboard-keys J and F are a special form of texture. Another variant of texture are Braille-letters carrying additional abstract information. Cluster ② has to be chosen when there is a need to present information on any surface via the tactile sense. This can be either coded information on a geological map on a more or less plane surface, but it can also be object specific features like the material itself. The resulting questions for the technical task are:

- Which body parts perform the interaction? → This trivial question has a significant impact, as the body part selected defines the resolution available on user side and consequently the requirements for the size of the texture-generating elements.
- Is the form of the texture-carrying shape subject to changes? If so, how much and in which areas? → If the shape changes a lot, it is likely that the unit providing the texture information has to be adapted to e.g. each finger (e.g. as a pin-array or piezoelectric disc), as the fingers will have to be positioned independently of each other. In this case it may even be necessary to provide a lateral movement

[4] Frequently the customer will not be able to specify these values directly. In this case creative variants of the question should be asked, e.g. by identifying the moving masses, or by taking measurements with one's own tools. These aids will be addressed within section 6.2.

between finger and texture-unit to generate shear forces in the skin. In case of the shape being fix, e.g. in the case of a map, a relative movement may happen by the fingers themselves and the texture unit can be designed with less size restrictions.
- How fast does the displayed information change? → Textures change rarely during the simulation of objects and display of maps. This is dramatically different when e.g. texts or the influences of fluids on textures have to be displayed. The answer to this question has a significant impact on the technical system.
- Which intensity range is covered by the texture? → In the simplest situation the answer can be given by definite displacements and a resolution in bit. Usually, only qualitative values of the properties of objects for interaction are available. These hints have to be complemented by one's own experiments. With regard to the definition of these requirements it is very important to make sure that the planned spatial addressability and maximum intensity change does not exceed the corresponding resolution of the user. A research on the corresponding psychophysical experiments is highly recommended, as otherwise it may not be possible to transmit the intended information density.

6.1.4 Cluster "Vibro-Tactile"

Cluster ③ is a solution space for simple one-dimensional technical problems and corresponding questions. It covers independent dimensions of information (e.g. coding an event in a frequency and the importance in the amplitude). In this cluster, distributions of intensity variations and /or time dependent distributions of single events are filed. Technological solutions are usually vibrational motors or tactons, as being used in mobile phones or game-consoles. But even if the technical solution itself seems quite straightforward, the challenge lies in the coding of information with respect to intensity and time and an appropriate mechanical coupling of the device to the user.

- Which mechanical interface for the transmission of haptic information to the user is planned? → More specifically: Is this interface influenced by mechanical limits like housings?
- Which design space is available? → Frequently, vibro-tactile solutions are limited as to the available space at an early stage of the design due to requirements for mobility.
- Which resolution is expected for the planned intensity variation? → The criterions are similar to those of the "surface-tactile" cluster. As the "vibro-tactile" cluster frequently deals with oscillating systems, the dependence of the perception of oscillations on its frequency has to be taken into account. The user's perception is the limiting factor for intensity variations, which themselves are dependent on the mechanical coupling between device and user, too.

6.1.5 Cluster "Vibro-Directional"

Vibro-tactile systems code one-dimensional information in the form of intensities. It is obvious that by the combination of multiples of such information sources directional information can be transmitted. This may happen two-dimensionally in a plane surface, but also three-dimensionally. Cluster ④ deals with such systems. One possible technical solution for directional surface information would be to locate a multitude of active units in the shape of a ring around a body part, e.g. like a belt around the belly. The direction is coded in the activity of single elements. This approach can also be transferred to a volumetric vector, whereby in these cases a large number of units is located on a closed surface, e.g. the upper part of the body. The activity of single elements codes the three dimensional direction as an origin of a normal vector on this surface. In addition to the questions of cluster ③ this cluster deals with the following questions:

- What is the intended resolution on the surface /in the space? → As well as before dependent on the body surface used, it is likely that the human perception represents the limit for the achievable resolution. Corresponding literature [51, 267] has to be checked carefully before the technical requirements can be met.
- What number of simultaneously displayed vectors is expected? → The fact that users will be able to identify one direction does not guarantee that with a parallel display of two points the user will perform equally well. Simultaneous display of information frequently results in masking-effects hard to be quantified. Experiments and analyses of the intended application are strongly recommended.
- Which frame of reference is used? → The information displayed is usually embedded in a frame of reference, which is not necessarily identical with the user's frame of reference and his or her body. The user may change his position for example in a vehicle, which results in a loss of the position of the elements fixed to the body and their orientation in the vehicle. It is necessary to be aware of the active frame of reference (local user-oriented, or vehicle-oriented, or maybe even world-oriented) and to provide measurement equipment for identifying changes in user positions and frame of reference. Additionally, it may become necessary to present a haptic reference signal to the user, which calibrates the user's perception to the frame of reference, e.g. a "north"-signal.

6.1.6 Cluster "Omni-Dimensional"

Cluster ⑤ deals with systems coding real volumetric information. Within such a three-dimensional space each point either includes intensity information (scalar field) or vector information (vector field). The sources of such data are numerous and frequent, may it be medical imaging data, or data of fluid mechanics, of atomic physics, of electrodynamics, or of electromagnetics. Pure systems of haptic inter-

action with such kinds of data are seldom. Frequently, they are combinations of the clusters “kinaesthetic” and “vibro-tactile” for scalar fields, respectively “kinaesthetic” with six active haptic degrees of freedom for vector fields[5]. Consequently, the specific questions of this cluster add one single aspect to already existing questions of the other clusters:

- Does the intended haptic interaction take place with scalar fields or with vector fields?→ For pure vector fields kinaesthetic systems with the corresponding questions for six active degrees of freedom should be considered. In the case of scalar fields, an analysis of vibro-tactile systems in combination with three-dimensional kinaesthetic systems and the corresponding questions should be considered. Then the property of the scalar value corresponds to the dynamics of the coded information.

6.1.7 Cluster “always”

For any development process there are several questions which always have to be asked. They often refer to the time-frame as well as to the resources available for the development. For haptic devices two specific questions have to be focused on, as they can become quite limiting for the design process due to specific properties of haptic devices:

- Which energy sources are available? → It is not a necessary prerequisite that electrical actuators have to be used for haptic devices, especially in the case of telemanipulation systems. The usage of pneumatic and hydraulic energy sources, especially for tactile devices is a real alternative and should be considered.
- The design, how expensive may it be? → The prices of current kinaesthetic haptic systems reach from 200 EUR of mass-products to 1.500 EUR of medium scale products to devices of 25t EUR for small series and. 100t EUR for individual solutions. These prices only partly result from commercial acquisitiveness, but mostly from the technical requirements and the efforts which have to be taken.

6.2 Experiments Together With the Customer

The customer formulates requirements - as mentioned before - typically in an inexact instead of a specific way. Additionally, there is the problem of a very unspecific terminology with regard to the design of haptic systems For the description of haptic sensual impressions there are numerous adjectives difficult to quantify, like: rough,

[5] The haptic interaction with objects in a mathematical abstraction always is an interaction with vector fields. In the vectors, forces of surfaces are coded, which themselves are time dependent, e.g. from movements and /or deformations of the objects themselves.

soft, smooth, gentle, mild, hard, viscous, as well as others derived from substantives such as: furry, silky, hairy, watery, and sticky which can be compared to real objects. So what could be more obvious than asking the customer to describe his / her haptic impressions by comparisons?

> Ask the customer to describe the intended haptic impression with reference to objects and items in his/her environment. These items should be easily at hand, like e.g. vegetables and fruits which offer a wide spectrum of textures and consistencies for comparison.

Sometimes the customer first needs to develop a certain understanding of the haptic properties of objects and items. This can best be achieved by his/her directly interacting with them. Examples of haptically extreme objects have to be included in a good sample case, too.

> Provide a sample case including weights and springs of different size, even marbles, fur, leather and silk. Depending on the project, add sandpaper of different granularity. Use these items to explain haptic parameters to the customer and help the customer to optimize the description of the product performance expected!

From practical experience, we can recommend also to take spring balances and letter balances or electronic force sensors with you to customer meetings. Frequently, it is possible to attach a handle directly to the items and ask the customer to pull, until a realistic force is reached. This enables customers of non-technical disciplines to quickly get an impression of the necessary moments and forces.

> Take mechanical measurement instruments with you to the customer meetings and allow the customer to touch and use them! This gives him / her a good first impression of the necessary force amplitudes.

In order to give a better impression of texture, mechanical workshops may produce patterns of knurls and grooves of different roughness on metals. Alternatively, sandpaper can be used and, by its defined grade of granularity, can provide a standardized scale to a certain extent.

> Use existing materials with scales to describe roughness and simulate the impression of texture.

What does not work:

A normal customer without expertise in the area of haptics will not be able to give statements concerning resolutions or dynamics of the haptic sense. This kind of information has to be derived from the type of interaction and the study of the psychophysical knowledge of comparable applications. Therefore, the experience of the developing engineer is still indispensable despite all the systematizations in the technical design process.

> Do not confuse the customer by asking questions about the physical resolution! This is necessarily the knowledge of the haptic engineer. However, learn about the dynamics of the interaction and try to assess the application, e.g by asking about the frame rate of a simulation, or the maximum number of load changes per seconds of a telemanipulator.

6.3 Requirement Specifications of a Haptic System

The customer ’s answers result in a functional requirement list. Depending on this list, a system specification sheet is formulated. This sheet will differ in its structures depending on whether tactile or kinaesthetic systems are addressed. In table 6.1 the most relevant technical parameters of a general haptic system are collected. This table is meant to be an orientation and has to be adapted to the specific situation by removing irrelevant points and adding application-specific aspects.

Table 6.1: Example of a system specification for a haptic device

SYSTEM SPECIFICATIONS OF A HAPTIC DEVICE Rev.: 01.01.2009			
R/W	DESCRIPTION	VALUE	SOURCE/COMMENT
R/W[6]	**Especially kinaesthetically motivated parameters**		
R	Number of DOFs	2x rot., 1x transl.	Shall give an idea of DOFs, name them!
R	Workspace	$100x50x50\,mm^3$	Minimum of workspace to be achieved
W	Maximum Workspace	$150x100x100\,mm^3$	Maximum workspace necessary
R	Maximum force in DOF "name"	5 N	
W	Maximum force in DOF "name"	7 N	Always define a range of forces!
R	Minimum force in DOF "name"	0.2 N	
W	Minimum force in DOF "name"	0.1 N	Always define a range of forces!
R	Maximum dynamics (bandwidth) for DOF "name" in a blocked situation	100 Hz	Shows (among other things) e.g. the maximum dynamics of the driver electronics
W[7]	Maximum dynamics (bandwidth) for DOF "name" in a blocked situation	200 Hz	Shows the bandwidth the customer dreams of.
R	Smallest border frequency when movement is blocked	static	There may be applications with pure dynamic movements without a static portion. This makes this question interesting.
R	Maximum velocity of movement in idle mode	10 mm/s	This is a question regarding security too, as it defines the mechanical energy stored in the system.
R	Maximum bandwidth of the velocity change	10 Hz	The change of velocity, which is the acceleration of the system, has a large influence on the energy the system requires.
R	Maximum haptic impedance at the output	10 Ns/m	This is an alternative representation to the independent definition of force and velocity for dynamic (but passive) systems!

Continuation on the next page ...

[6] R: requirement, W: wish

[7] The combination of requirements and wishes (R and F) may be used for almost any element of the system specification. It is recommended to make use of this method, but due to clarity in the context of this book this approach of double-questions is aborted here.

Continuation:

Rt/W	Description	Value	Source/Comment
R	Minimum haptic impedance at the output	0.01 Ns/m	This is an alternative representation to the independent definition of force and velocity for dynamic (but passive) systems!
R	Smallest position resolution/measurement insecurity for DOF "name"	0.1 mm	Usually measurement of the position is self-evident for haptic interaction.
W	Smallest position resolution/measurement insecurity for DOF "name"	0.05 mm	
R	Type of the mechanical interface	Button/pen/none	Is there a handle?
R	Mechanical reference point	Grounded, worn	Has influence on weight, size and energy
R/W	**Especially tactile motivated parameters**		
R	Direction(s) of the tactile stimulation	Normal to the skin	An alternative would be laterial or a combinationof both
R	Maximum displacement-amplitude of the tactile elements	1 mm	Is especially relevant for pin-displays, but may be also understood as oscillation-amplitude of vibrational elements.
R	Minimum amplitude resolution of displacements	digital (on/off)	May include several levels for the pin to be moved to
R	Highes density of stimulation	2 mm Distance from midpoint ot midpoint	Varies extremely in dependency from the chosen skin area in contact
R	Maximum geometrical size of stimulation	2 mm Diameter	
R	Maximum frequency range of stimulation	100 to 300 Hz	Relevant for tactile actuators only, of course.
R	Minimum frequency-resolution	1 Hz	For vibrotactile actuators
R	Maximum force during displacement/stiffness	20 N	Pin-based actuators may not necessarily be stiff. Systems of lower admittance may be used too.
R	Connection to the user	Attached to the environment / worn	Necessary to identify, whether there is a relative movement between skin (e.g. finger) and the display.
R	Maximum number of fingers simultaneously in contact with the device	1-10	May have an immensive impact on the design when e.b. full-hand exploration is required.
R/W	**Digital interface**		

Continuation on the next page ...

Continuation:

RT/W	DESCRIPTION	VALUE	SOURCE/COMMENT
R	Minimum resolution of the ouput data	12 bit	Usually slightly lower than the measurement error of force- and position-measurement.
R	Minimum resolution of the input data	12 bit	Usually slightly larger than the resolution of force- and position input-data.
R	Frequency of the haptic loop[8]	1000 Hz	Should be at least two times, better would be 10 times, larger than the border frequency of the design. Has influence on the perceived stiffness.
W	Other interface-requirements	Use USB/FireWire...	Typically the interface to be used is subject to company politics.
R	Interface driver	API	As any other hardware a haptic interface needs an own software driver for abstraction.
R/W	**General parameters**		
R	Maximum temperature range for operation	10°C to 50°C	May become very relevant for actuator principals with little efficiency in extreme environments (automotive).
R	Maximum volume	$500 \cdot 500 \cdot 200\,\text{mm}^3$	Device-size
R	Weight	1 kg	Especially relevant if the device is worn. This limit will strongly influence the mechanical energy generated.
R	Electrical supply	Battery/ 110V/ 230V	Very important, devices were spotted on fairs, which ceased to function due to errors made when considering AC voltages of different countries.
R	Maximum power	50 W	Primary power consumption including all losses.
R	Additionally a system specification includes references to other standards and special requirements relating to the product development process. Among others, these are the costs for the individual device, the design-process itself and the number of devices to be manufactured in a certain time-frame. Additionally the time of shipment, visual parameters for the design, and security related issues are usually adressed.		

[8] A “haptic loop” is a complete cylcle including the output of the control-variable (in case of simulators this variable was calculated the time-step before) and the read operation on the measurement value

6.4 Order of Technical Decisions

During the technical development of haptic systems, the decisions on single components influence each other intensively. However, this influence is not identical between all components. For the engineer it is necessary to proceed in the solution identification for each component, after having gained the knowledge of the requirements for the haptic system. It is obvious that, according to a systematic development process, each solution has to be compared to the specifications concerning its advantages and disadvantages. The recommended procedure of how to deal with the components is the basis of the chapter structure of this section of the book and is summarized once again for completeness:

1. Decision about the control engineering structure of the haptic system based on the evaluation of the application (tactile or kinaesthetic), the impedance in idle state (masses >20 g and friction acceptable) and the maximum impedance (stiffnesses > 300 N/m or smaller) (fig. 6.2, chapter 7)
2. Decision about the kinematics based on calculations of the working volume and the expected stiffness (chapter 8)
3. Based on the well-known mechanical structure, the actuator design can be made. For this a rough decision can be taken, based on section 9.1 for orientation, followed by the concrete dimensioning and the calculations of the actuators according to the other sections in chapter 9.1.
4. Dependent on the chosen control engineering structure, the force-sensor design can be performed parallel to the actuator design (chapter 10).
5. Relatively uncritical for the design is the choice of the positioning sensors (chapter 11).
6. The electronic interfaces are subordinate to all the decisions made before (chapter 12).
7. The software design of the haptic rendering itself ,in many aspects, is so independent of prior technical decisions that it can be decoupled in almost all aspects (chapter 13) from the rest of the design, - however not of the common set of specifications, of course.

Nevertheless it is vital to note that e.g. the kinematics design cannot be realized, completely decoupled from the available space for the device and the forces and torques - respectively the actuator. Additionally, kinematics directly influences any measurement technology as even displacement sensors have limitations on resolution and dynamics. The order suggested above for taking decisions has to be understood as being a recommendation for processing the tasks; it does not free the design engineer from the responsibility to keep an overview of the subcomponents and their reciprocal influence.

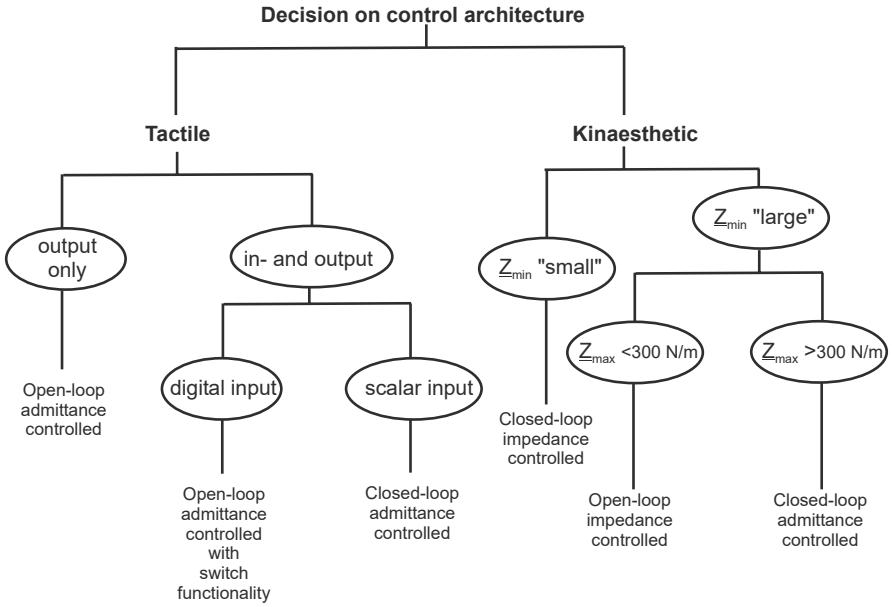


Fig. 6.2 Aid for the decision on the choice of the control structure.

"*This page left intentionally blank.*"

Chapter 7
Control of Haptic Systems

OLIVER MECKEL

7.1 Introduction

The analysis and investigation of technical systems and processes is based upon the requirement to establish safe and reliable system behavior, and controllable system states. By the depiction as *systems* the analysis is put on an abstracted level which allows covering many different technical systems described by their fundamental physics. On this abstracted level a general analysis of closed loop control issues is possible using several methods and techniques. The resulting procedures are applicable to a various number of system classes. The main purpose of any depiction and analysis of control systems is to achieve safe system behavior and reliable processes. *Power is nothing without control*, that might be a credo and the fundamental motivation for any system investigation to prove system reliability and safe system behavior.

The abstract description that shall be used for a closed loop control analysis starts with the mathematical formulation of the physical principles the system follows. As mentioned above, systems with different physical principles are covered up by similar mathematical methods. The depiction by differential equations or systems of differential equations proves widely usable for the formulation of various system behaviors. Herein analogies allow transforming this system behavior into the different technical context of a different system, provided that there exists a definite formulation of the system states that are of interest for the closed loop control analysis. The mathematical formulation of the physical principles of the system, also denoted as modeling, is followed by the system analysis including dynamic behavior and its characteristic. With this knowledge, a wide variety of design methods for control systems in closed loops become applicable. Their main requirements are:

T.A. Kern (ed.), *Engineering Haptic Devices*, DOI 10.1007/978-3-540-88248-0_7,
© Springer-Verlag Berlin Heidelberg 2009

- System stability: The fundamental requirement for stability in any technical system is the main purpose for closed loop control design.
- Control quality: Tracking behavior of the system states to demanded values every system is faced with external influences also denoted as disturbances which interfere with the demanded system inputs and disrupt the optimal system behavior. To compensate this negative influence a control system is designed.
- Dynamic behavior and performance: In addition to the first two issues, the need for a certain system dynamic completes this requirements-list. Besides the quality of the control result tracking the demanded values, the system behavior within the range of changes from these demanded values is focused. Also the control effort which needs to achieve a certain control result is to be investigated.

The main challenge for closed loop control design in many other engineering disciplines is to deal with different goals, that are often in conflict with each other. These conflicts are a direct result of the above characteristics. Often these issues are weighted in a different way depending on the used design method. Typically a gained solution is never an optimal one, rather a trade off between system requirements.

7.2 System Description

A variety of description methods can be applied for the mathematical formulation of systems with different physical principles. One of the main distinctions is drawn between methods for the description of linear and nonlinear systems, summarized in the following paragraphs.

7.2.1 Linear SISO-System Description

As shown in chapter 5, a formulation of haptic devices as impedance controlled respectively admittance controlled system leads to a certain specific definition of system states. Within both methods of formulation, the physical principles can be described by the corresponding differential equations. As a simple example a second order system is analyzed to show the similarity of both approaches. Figure 7.1 shows the model scheme of a damped mechanical oscillator with the parameters mass m, proportional damping d and linear stiffness k. The basic mechanical principles form the following differential equation:

$$m\ddot{x}(t) + d\dot{x}(t) + kx(t) = f(t) \tag{7.1}$$

Here x denotes the actual position of the mass m. Using this formulation the speed follows from the gradient $\frac{dx}{dt} = v$. The corresponding derivative of the differential equation leads to

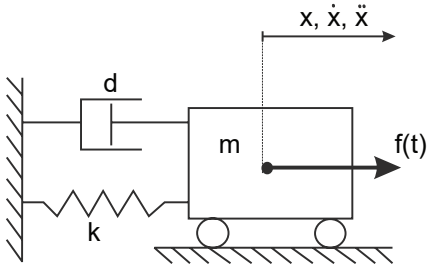


Fig. 7.1 Mechanical oscillator.

$$m\dot{v}(t) + dv(t) + k\int v\,\mathrm{d}t = f(t) \tag{7.2}$$

Looking on the electric RLC-system shown in figure 7.2, a similar formulation by a linear differential equation is formulated.

$$L\frac{di(t)}{dt} + Ri(t) + \frac{1}{C}\int i\,\mathrm{d}t = u(t) \tag{7.3}$$

The correspondence of these two differential equations relates the respective kinetic and potential energy storing mechanical devices – mass and spring – to the electrical energy storing devices – inductance and capacitance. Furthermore the dissipative mechanical damping is related to the electrical resistance. Another approach of system formulation leads to a second way how to relate mechanical and electrical system devices to each other as shown in the following differential equation:

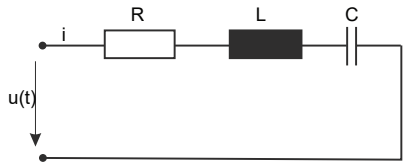


Fig. 7.2 RLC-Network.

$$C\frac{du(t)}{dt} + \frac{1}{R}u(t) + \frac{1}{L}\int u\,\mathrm{d}t = i(t) \tag{7.4}$$

Deriving from this system formulation the electric impedance $\underline{Z} = \frac{\underline{u}}{\underline{i}}$ describes the system transfer behavior. The reciprocal value denotes the electrical admittance $\underline{Y} = \frac{1}{\underline{Z}}$. This impedance and admittance respectively based system formulation is

transferable from the electrical RLC-systems to mechanical systems. Appendix 16 shows the general impedance or admittance based formulation of translatory and rotary mechanical systems. The main distinction is based on the definition of the input and output variables. For mechanical impedances the system input will be depicted by a certain speed or differential speed variable, the system output variable will denote a resulting force. The formulation of a mechanical admittance switches input and output variable which leads to a system input describing a force affecting the energy storing devices and influencing the speed as the system output. In many approaches the description of mechanical systems is based on the equilibrium of forces, which often leads to formulation of the system behavior as mechanical admittance. Nevertheless the formulation as mechanical impedance is not a mistake; it is merely a change of the point of view on the same system principle. The following example shows both approaches:

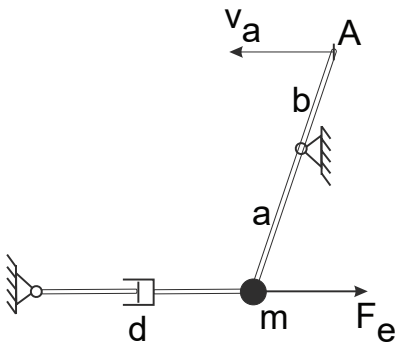


Fig. 7.3 Example for mechanical admittance.

Within this system F_e denotes the force affecting the mass m as the system input. As output variable the system speed in point A v_a is chosen. The system behavior is described by the following linear differential equation:

$$m\dot{v}_m + dv_m = F_e \tag{7.5}$$

The kinematic of the links provides $v_m = \frac{b}{a}v_a$ which leads to

$$m\frac{a}{b}\dot{v}_a + d\frac{a}{b}v_a = F_e. \tag{7.6}$$

Switching the point of view and choosing a measurable force in point A as system output and the speed of the mass $v_m = v_e$ as system input as shown in figure 7.4

The resulting differential equation is

$$F_a = m\frac{a}{b}\dot{v}_e + d\frac{a}{b}v_e. \tag{7.7}$$

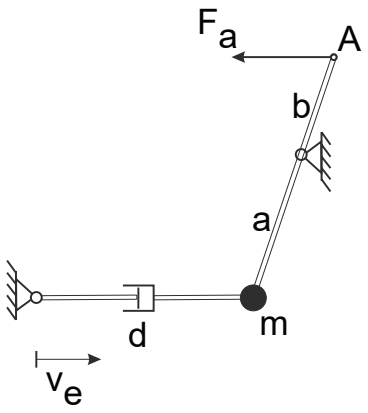


Fig. 7.4 Example for mechanical impedance.

The transformation of impedances or admittances into the LAPLACE domain abstracts the system analysis, and a flexible investigation of coupled system components becomes possible. In many control design methods the LAPLACE domain is used to gain a consistent formulation of system behavior described as LAPLACE transfer function $\underline{G}(s)$. Especially while dealing with Single-Input-Single-Output Systems this method of description has proven its strength and allows the direct applicability of a various number of techniques for stability analysis, system classification and control system design procedures. Herein the distinction between impedance and admittance based system structures is not necessary any more. The transformation to the LAPLACE domain of both differential equations from the example above brings up the following transfer functions

$$\frac{\underline{F_a}(s)}{\underline{v_e}(s)} = \frac{a}{b} \cdot (d + m \cdot s) \tag{7.8}$$

$$\frac{\underline{v_a}(s)}{\underline{F_e}(s)} = \frac{b}{a} \cdot \frac{1}{d + m \cdot s}. \tag{7.9}$$

It becomes quite obvious that one of these transfer function is the reciprocal of the other. The corresponding phase plots point out the main difference between both formulations. Describing the exemplary system as admittance as described in equation 7.6, $\underline{v_a}(s)$ as output signal shows a phase delay to the input signal $\underline{F_e}(s)$, whereas the description of the system using impedances leads to exactly the opposite behavior. For comparison the step responses of both formulations are investigated.

Due to the different choice of input and output variables the system shows a different transfer behavior. This behavior described by equation 7.9 shows as expected the typical step response of a PT_1-System whereas the formulation of equation 7.8 leads to a DT_1 -behavior.

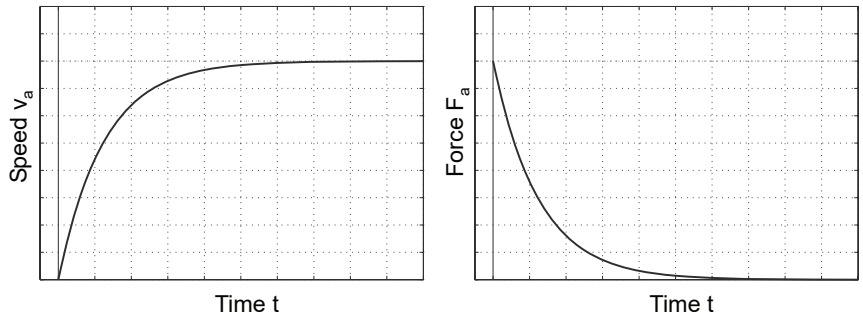


Fig. 7.5 Step response for system analysis.

In figure 7.5 on the left side the PT_1 step response in the speed v_a due to a force F_e is shown. The right side describes the step response of the force F_a due to a speed v_e. The differences of these step responses become quite obvious, each corresponding to the based system formulation, and can be justified by the different particular solution of the differential equation.

The shown methods to establish a mathematical formulation of a system are completely equivalent. Consequently the use of LAPLACE transfer function $\underline{G_i}(s)$ permits a nearly arbitrary combination of different system components, which can be independently modelled as impedance or admittance. For further information about the theoretical background on the LAPLACE domain and the corresponding transformation, [264], [60] and [165] are recommended.

7.2.2 Linear State Space Description

Besides the formulation of system characteristics through transfer functions, the description of systems using the state space representation in the time domain allows to deal with arbitrary linear systems too. For Single-Input-Single-Output-Systems, a description using an *n*th order ordinary differential equation is transformable into a set with *n* first order ordinary differential equations. In addition to the simplified usage of numerical algorithms for solving this set of differential equations, the major advantage is the applicability to Multi-Input-Multi-Output-Systems. A correct and systematic model of their coupled system inputs, system states, and system outputs is comparably easy to achieve. On the contrary to the system description in the LAPLACE domain by transfer functions $\underline{G}(s)$, the state space representation formulates the system behavior in the time domain. Two set of equations are necessary for a complete state space system representation. These are denoted as the *system equation*

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} \tag{7.10}$$

and the *output equation*

$$\mathbf{y} = \mathbf{C}\mathbf{x} + \mathbf{D}\mathbf{u}. \tag{7.11}$$

The vectors **u** and **y** describe the multidimensional system input respectively system output. Vector **x** denotes the inner system states.

As an example for state space representation the 2nd order mechanical oscillating system as shown in figure 7.6 is examined. Assuming the existence of time invariant parameters the description by a 2nd order differential equation is:

$$m\ddot{y} + d\dot{y} + ky = u \tag{7.12}$$

Figure 7.7 shows the corresponding block diagram of this PT_2-System.

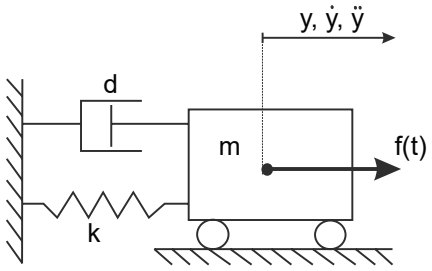


Fig. 7.6 2nd order oscillator.

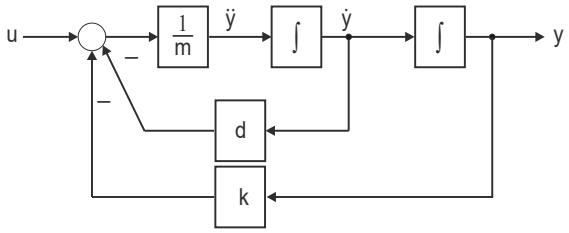


Fig. 7.7 PT_2 block diagram.

The transformation of the 2nd order differential equation 7.12 into a set of two 1st order differential equation is done by choosing the integrator outputs as system states:

$$x_1 = y \Rightarrow \dot{x}_1 = x_2$$
$$x_2 = \dot{y} \Rightarrow \dot{x}_2 = -\frac{k}{m}x_1 - \frac{d}{m}x_2 + \frac{1}{m}u \tag{7.13}$$

Thus the system equation for the state space representation is as follows:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{d}{m} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix} u \tag{7.14}$$

The general form of the system equation is:

$$\dot{\mathbf{x}} = \mathbf{A}\,\mathbf{x} + \mathbf{B}\,\mathbf{u} \tag{7.15}$$

This set of equations contains the *state space vector* **x**. Its components describe all inner variables of the process that are of interest and that have not been examined explicitly using a formulation by transfer function. The system output is described by the output equation. In the given example as shown in figure 7.6 the system output y is equal to the inner state x_1

$$y = x_1 \tag{7.16}$$

which leads to the vector representation of

$$y = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{7.17}$$

The general form of the output equation is:

$$\mathbf{y} = \mathbf{C}\,\mathbf{x} + \mathbf{D}\,\mathbf{u} \tag{7.18}$$

which leads to the general state space representation that is applicable for Single or Multi Input and Output systems. The structure of this representation is depicted in figure 7.8:

$$\dot{\mathbf{x}} = \mathbf{A}\,\mathbf{x} + \mathbf{B}\,\mathbf{u}$$
$$\mathbf{y} = \mathbf{C}\,\mathbf{x} + \mathbf{D}\,\mathbf{u}$$

Although not mentioned in this example, matrix **D** denotes a direct feedthrough which occurs in systems whose output signals **y** are directly affected by the input signals **u** without any time delay. Thus these systems show a non delayed step response. For further explanation on $\mathbf{A}, \mathbf{B}, \mathbf{C}$ und **D** the reference [164] is recommended.

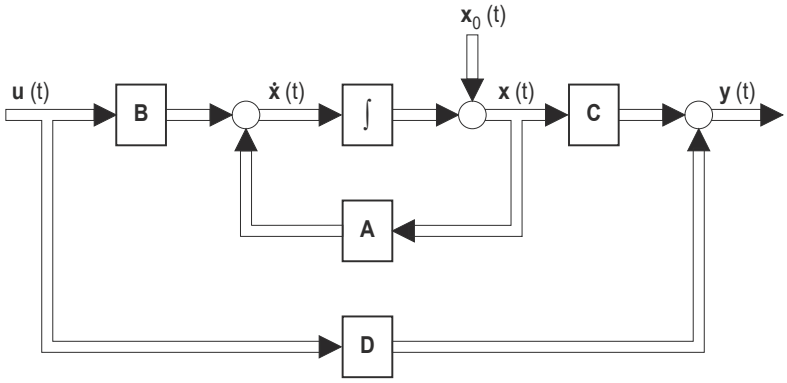


Fig. 7.8 State space description.

7.2.3 Nonlinear System Description

A further challenge within the formulation of system behavior is to imply nonlinear effects, especially if a subsequent system analysis and classification is needed. Although a mathematical description of nonlinear system behavior might be found fast, the applicability of certain control design methods is an additional problem. Static nonlinearities can be easily described by a serial coupling of a static nonlinearity and linear dynamic device to be used as a summarized element for closed loop analysis. Herein two different models are differentiated. Figure 7.9 shows the block diagram consisting of a linear element with arbitrary subsystem dynamics followed by a static nonlinearity. This configuration also known as WIENER-*Model* is

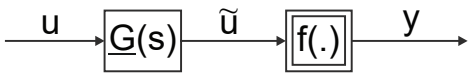


Fig. 7.9 WIENER-Modell.

described by

$$\tilde{u}(s) = \underline{G}(s) \cdot u(s)$$
$$y(s) = f(\tilde{u}(s)).$$

In comparison, figure 7.10 shows the configuration of the HAMMERSTEIN-*Model* changing the order of the underlying static nonlinearity and the linear dynamic subsystem. The corresponding mathematical formulation of this model is as described

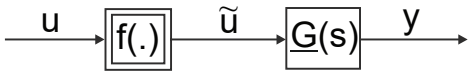


Fig. 7.10 HAMMERSTEIN-Modell.

by

$$\tilde{u}(s) = f(u(s))$$
$$y(s) = \underline{G}(s) \cdot \tilde{u}(s).$$

More complex structures appear as soon as the dynamic behavior of a system is affected by nonlinearities. Figure 7.11 shows as an example a system with an internal saturation of a states derivative. For this configuration both models cannot be applied as easily as for static nonlinearities furthermore if a system description is needed usable for certain methods of system analysis and investigation.

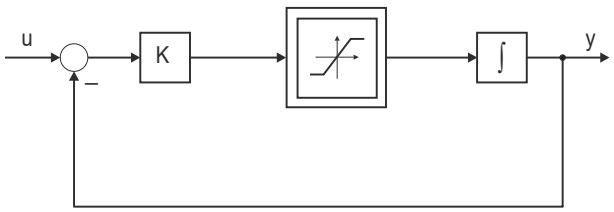


Fig. 7.11 System with internal saturation.

Typical examples for systems showing that kind of nonlinear behavior are electrical motors whose torque current characteristic is affected by saturation effects, and thus whose torque available for acceleration is limited to a maximum value.

This kind of system behavior is one example of how complicated the process of system modeling may become, as ordinary linear system description methods are not applicable to such a case. Nevertheless it is necessary to gain a system formulation in which the system behavior and the system stability can be investigated

successfully. To achieve a system description taking various system nonlinearities into account, it is recommended to set up a nonlinear state space descriptions. They offer a wide set of tools applicable to the following investigations. Deriving from equation 7.10 and 7.11 the nonlinear system description for single respectively multi input and output systems is as follows:

$$\dot{\mathbf{x}} = \mathbf{f}(x, \mathbf{u}, t)$$
$$\mathbf{y} = \mathbf{g}(x, \mathbf{u}, t).$$

This state space description is most flexible to gain a usable mathematical formulation of a systems behavior consisting of static, dynamic and arbitrarily coupled nonlinearities.

7.3 System Stability

As mentioned in section 7.1 one of the most important goals of the control design is the stabilization of systems or processes during their life cycle, while operative or disabled. Due to the close coupling of haptic systems to a human user via a human machine interface, safety becomes most relevant. Consequently the focus of this chapter lies on system stability and its analysis by using certain methods applicable to many systems. It has to resemble the system's behavior correctly, and has to be aligned with applied investigation technique. For the investigation of systems, subsystems, closed looped systems, and single or multi input output systems, a wide variety of different methods exists. The most important ones shall be introduced in this chapter.

7.3.1 Analysis of Linear System Stability

The stability analysis of linear time invariant systems is easily done by the investigation of the system poles or roots derived from the eigenvalue calculation of the *system transfer function* $\underline{G}(s)$. The decisive factor is the sign of the real part of these system poles. A negative sign in this real part indicates a stable eigenvalue, a positive sign denotes an unstable eigenvalue. The correspondence to the system stability becomes obvious while looking at the homogenous part of the solution of the ordinary differential equation describing the system behavior. As example a system shall be described by

$$T\dot{y}(t) + y(t) = Ku(t). \tag{7.19}$$

The homogenous part of the solution $y(t)$ is derived using

$$y_h = e^{\lambda t} \quad \text{with} \quad \lambda = -\frac{1}{T}. \tag{7.20}$$

As it can be seen quite fast the pole $\lambda = -\frac{1}{T}$ has a negative sign only if the time constant T has a positive sign. In this case the homogenous part of $y(t)$ disappears for $t \to \infty$, while it rises beyond each limit exponentially if the pole $\lambda = -\frac{1}{T}$ is unstable. This section will not deal with the basic theoretical background of linear system stability, as these are basics of control theory. Focus of this section is the application of certain stability analysis methods. Herein it will be distinguished between methods for a direct stability analysis of a system or subsystem and techniques of the closed looped stability analysis. For direct stability analysis of linear system the investigation of the poles placement in the complex plane is fundamental. Besides the explicit calculation of the system poles or eigenvalues the ROUTH HURWITZ *criterion* offers to determine the system stability and the system pole placement with explicit calculation. In many cases this simplifies the stability analysis. For the analysis of the closed loop stability the determination of the closed loop pole placement is also a possible approach. Additional methods leave room for further design aspects and extend the basic stability analysis. Well-known examples of such techniques are

- Root locus method
- NYQUIST´s stability criterion.

The applicability of both methods will be discussed in the following without looking at the exact derivation. The root locus offers the opportunity to investigate the pole placement in the complex plane depending on certain invariant system parameters. As example of invariant system parameters changing time constants or variable system gains might occur. The gain of the open loop is often of interest within the root locus method for closed loop stability analysis and control design. In equation 7.21 G_R denotes the transfer function of the controller, G_S describes the behavior of the system, plant or process to be controlled.

$$-G_o = G_R G_S \tag{7.21}$$

Using the root locus method, it is possible to apply predefined sketching rules whenever the dependency of the closed loop pole placement on the open loop gain K is of interest. The closed loop transfer function is depicted by equation 7.22

$$G_g = \frac{G_R G_S}{1 + G_R G_S} \tag{7.22}$$

As an example an integrator system with a second order delay (IT_2) described by

$$G_S = \frac{1}{s} \cdot \frac{1}{1+s} \cdot \frac{1}{1+4s} \tag{7.23}$$

is examined. The control transfer function is $G_R = K_R$. Thus we find as open loop transfer function

$$-G_o = G_R G_S = \frac{K_R}{s(1+s)(1+4s)}. \tag{7.24}$$

Using the sketching rules which can be found in various examples in literature [165], [264], the root locus graph as shown in figure 7.12 is derived. The graph indicates,

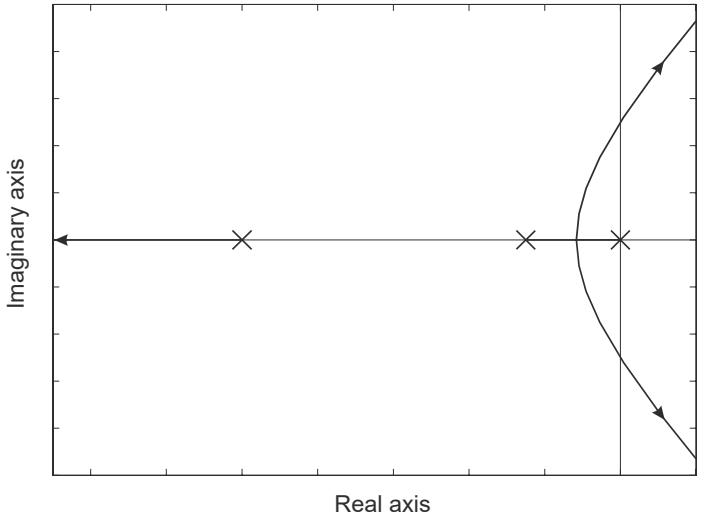


Fig. 7.12 IT_2 root locus.

that small gains K_R lead to a stable closed loop system since all roots have a negative real part. A rising K_R leads to two of the roots crossing the imaginary axis and the closed loop system becomes unstable. This simplified example proves that this method can easily be integrated in a control design process, as it delivers a stability analysis of the closed loop system only processing an examination on the open loop system. This issue is also one of the advantages of the NYQUIST stability criterion. Additionally the definition of the open loop system is sufficient to derive a stability analysis of the system in a closed loop arrangement. This section will concentrate on the simplified NYQUIST stability criterion investigating the open loop frequency response described by

$$-G_o(j\omega) = G_R(j\omega)G_S(j\omega). \tag{7.25}$$

The NYQUIST stability criterion is based on the characteristic correspondence of amplitude and phase of the frequency response. As example we use the already introduced IT_2-system controlled by a proportional controller $G_R = K_R$. The Bode plot of the frequency response is shown in figure 7.13:

The stability condition which has to be met is given by the phase of the open loop frequency response, with $\varphi(\omega) > -180°$ in case of the frequency response's amplitude $A(\omega)$ being above 0 dB. As shown in figure 7.13, the choice of the controller

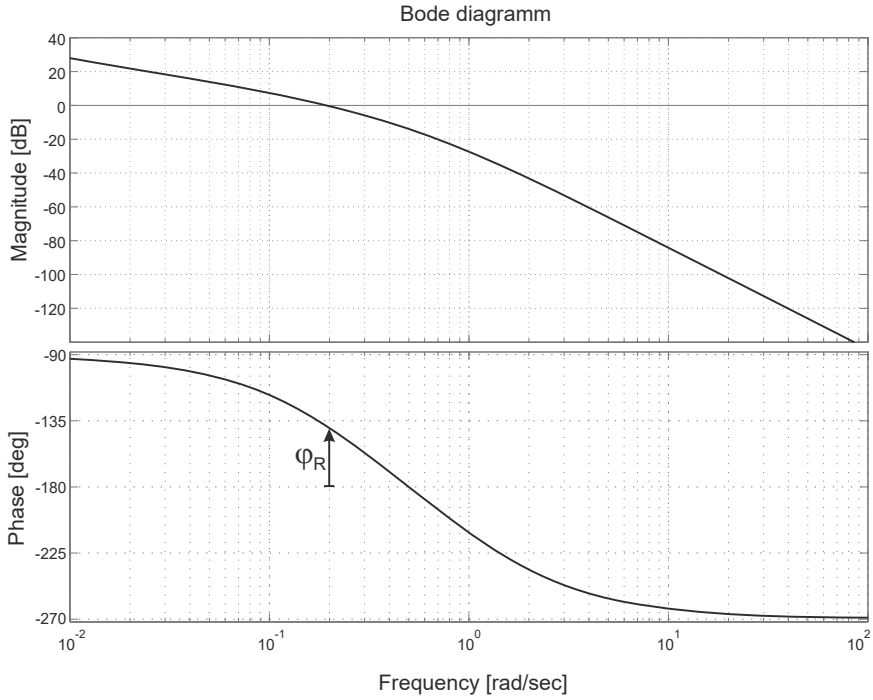


Fig. 7.13 IT_2 frequency response.

gain K_R transfers the amplitude graph of the open loop frequency response vertically without affecting the phase of the open loop frequency response. For most applications the specific requirement of a sufficient phase margin φ_R is compulsory. The resulting phase margin is also shown in figure 7.13. All such requirements have to be met in the closed control loop and must be determined to choose the correct control design method. In this simplified example the examined amplitude and phase of the open loop frequency response is dependent on the proportional controller gain K_R, which is sufficient to establish system stability including a certain phase margin. More complex control structures such as PI, $PIDT_n$ or Lead Lag extend the possibilities for control design to meet further requirements.

This section showed the basic principle of the simplified NYQUIST criterion being applicable to stable open loop systems. For an investigation of unstable open loop systems the general form of the NYQUIST criterion must be used, which itself is not introduced in this book. For this basic knowledge it is recommended to consult [165] and [264].

7.3.2 Analysis of Nonlinear System Stability

The application of all previous approaches for the analysis of system stability is limited to linear time invariant systems. Nearly all real systems show nonlinear effects or consist of nonlinear subsystems. One approach to deal with these nonlinear systems is the linearization in a fixed working point. All further investigations are focused on this point, and the application of the previously presented methods becomes possible. If these methods are not sufficient, extended techniques for stability analysis of nonlinear systems must be applied. The following are examples of representing completely different approaches:

- Principle of the harmonic balance
- Phase plane analysis
- POPOV criterion and circle criterion
- LYAPUNOV´s direct method
- System passivity analysis

Without dealing with the mathematical background or the exact proof the principles and the application of chosen techniques shall be demonstrated. At this point a complete explanation of this topic is too extensive due to the wide variety of the underlying methods. For further detailed explanation study [58], [56], [57], [265], [135] and [234].

As an introductionary example the analysis of closed loop systems can be done applying the POPOV criterion respectively the circle criterion. Figure 7.14 shows the block diagram of the corresponding closed loop structure of the system that is going to be analyzed:

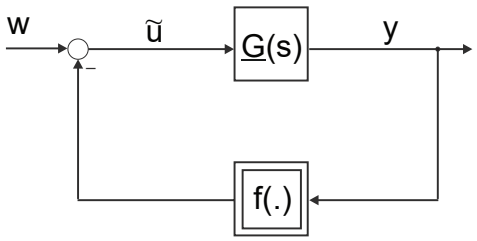


Fig. 7.14 Nonlinear closed loop system.

The bock diagram consists of a linear transfer function $\underline{G}(s)$ with arbitrary dynamics and a static nonlinearity $f(.)$. The state space formulation of $\underline{G}(s)$ is as follows:

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\tilde{u}$$
$$\mathbf{y} = \mathbf{C}\mathbf{x}$$

Thus we find for the closed loop system description:

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} - \mathbf{B}f(y)$$
$$y = \mathbf{C}\mathbf{x}.$$

In case that $f(y) = k \cdot y$ this nonlinear system is reduced to a linear system whose stability can be examined with the evaluation of the system's eigenvalues. With an arbitrary nonlinear function f(y) the complexity of the problem is extended. So first constraint to f(y) is that it exists only in a determined sector that is limited by a straight line through the origin with a gradient k. Figure 7.15 shows an equivalent example for the nonlinear function $f(y)$. This constraint is depicted by the following equation:

$$0 \leq f(y) \leq ky. \tag{7.26}$$

The POPOV criterion provides an intuitive handling for the stability analysis of the presented example. The system is asymptotically idle state ($\dot{\mathbf{x}} = \mathbf{x} = \mathbf{0}$) stable if:

- the linear subsystem $\underline{G}(s)$ is asymptotically stable and fully controllable,
- the nonlinear function meets the presented sector condition as shown in figure 7.15,
- for an arbitrarily small number $\rho \geq 0$ there exists a positive number α, so that the following inequality is satisfied:

$$\forall \omega \geq 0 \quad \mathrm{Re}[(1 + j\alpha\omega)\underline{G}(j\omega)] + \frac{1}{k} \geq \rho \tag{7.27}$$

Equation 7.27 formulates the condition also known as POPOV *inequality*. With

$$\underline{G}(j\omega) = \mathrm{Re}(\underline{G}(j\omega)) + j\mathrm{Im}(\underline{G}(j\omega)) \tag{7.28}$$

equation 7.27 leads to

$$\mathrm{Re}(\underline{G}(j\omega)) - \alpha\omega\mathrm{Im}(\underline{G}(j\omega)) + \frac{1}{k} \geq \rho \tag{7.29}$$

With an additional definition of a related transfer function

$$G^* = \mathrm{Re}(\underline{G}(j\omega)) + j\omega\mathrm{Im}(\underline{G}(j\omega)) \tag{7.30}$$

equation 7.29 states that the plot in the complex plane of $\underline{G}^*$, the so called POPOV *plot*, has to be located in a sector with an upper limit described by $y = \frac{1}{\alpha}(x + \frac{1}{k})$. Figure 7.16 shows an example for the POPOV plot of a system in the complex plane constrained by the sector condition.

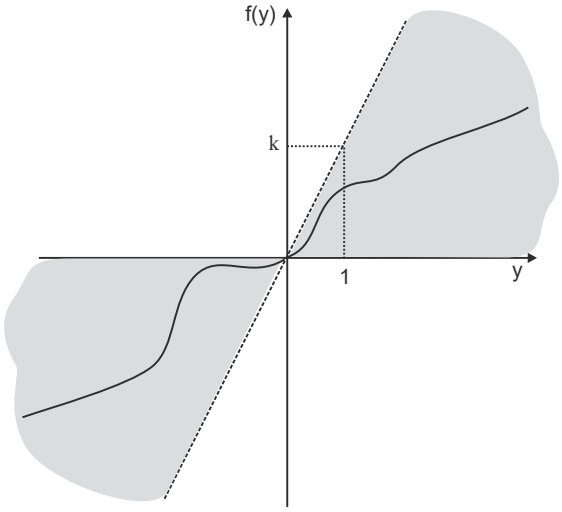


Fig. 7.15 Sector condition.

The close relation to the NYQUIST criterion for the stability analysis of linear systems becomes quite obvious here. While the NYQUIST criterion examines the

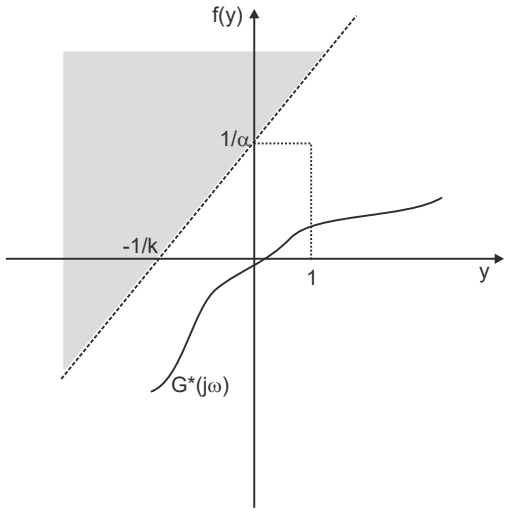


Fig. 7.16 POPOV plot.

plot of $\underline{G}(j\omega)$ referred to the critical point (-1|0), the location of the POPOV plot is checked for a sector condition defined by a straight line limit.

The application of the POPOV criterion has the excelling advantage, that it is possible to gain a result out of the stability analysis without an exact formulation of the nonlinearity within the system. All constraints for the nonlinear subsystem are restraint to the sector condition and the condition to have memoryless transfer behavior. The most complicated aspect within this kind of analysis is how to formulate the considered system structure in a way, that the POPOV criterion can be applied. For completeness the circle criterion shall be mentioned whose sector condition is not represented by a straight line, rather

$$k_1 \leq \frac{f(y)}{y} \leq k_2. \tag{7.31}$$

defines the new sector condition. For additional explanation on these constraints and the application of the circle criterion it is recommended to consider [265], [135] and [234].

As second example for stability analysis of nonlinear systems the *direct method by* LYAPUNOV is introduced. The basic principle is that if both linear and nonlinear stable systems tend to a stable steady state, the complete system energy has to be dissipated continuously. Thus it is possible to gain result from stability analysis while verifying the characteristics of the function representing the state of energy in the system. LYAPUNOV´s direct method generalizes this approach to evaluate the system energy by the generation of an artificial scalar function which can describe not only the energy stored within the considered dynamic system, further it is used as an energy like function of a dissipative system. These kinds of functions are called LYAPUNOV *functions* $V(\mathbf{x})$.

For the examination of the system stability the already mentioned state space description of a nonlinear system is used:

$$\begin{aligned} \dot{\mathrm{x}} &= \mathrm{f}(\mathrm{x},\mathrm{u},t) \\ \mathrm{y} &= \mathrm{g}(\mathrm{x},\mathrm{u},t). \end{aligned}$$

By the definition of LYAPUNOV´s theorem the equilibrium at the phase plane origin $\dot{\mathbf{x}} = \mathbf{x} = \mathbf{0}$ is globally, asymptotically stable if

1. a positive definite scalar function $V(\mathbf{x})$ with $\mathbf{x}$ as the system state vector exists, meaning that $V(\mathbf{0}) = 0$ and $V(\mathbf{x}) > 0 \forall \mathbf{x} \neq \mathbf{0}$,
2. $\dot{V}$ is negative definite, meaning $\dot{V}(\mathbf{x}) \leq 0$,
3. $V(\mathbf{x})$ is not limited, meaning $V(\mathbf{x}) \rightarrow \infty$ as $\| x \| \rightarrow \infty$.

If these conditions are met in a bounded area at the origin only, the system is locally asymptotically stable.

As a clarifying example the following nonlinear first order system

$$\dot{x} + f(x) = 0 \tag{7.32}$$

is evaluated. Herein $f(x)$ denotes any continuous function of the same sign as its scalar argument x so that $x \cdot f(x) > 0$ and $f(0) = 0$. Applying this constraints a LYAPUNOV function candidate can be found described by

$$V = x^2. \tag{7.33}$$

The time derivative of $V(x)$ provides

$$\dot{V} = 2x\dot{x} = -2xf(x). \tag{7.34}$$

Due to the assumed characteristics of $f(x)$ all conditions of LYAPUNOV´s direct method are satisfied thus the system has globally asymptotically stable equilibrium at the origin. Although the exact function $f(x)$ is not known, the fact that it exists in the first and third quadrant only is sufficient for $\dot{V}(x)$ to be negative definite. As second example a multi-input multi-output system is examined depicted by its state space formulation

$$\begin{aligned} \dot{x}_1 &= x_2 - x_1(x_1^2 + x_2^2) \\ \dot{x}_2 &= -x_1 - x_2(x_1^2 + x_2^2). \end{aligned}$$

In this example the system has an equilibrium at the origin too. Consequently the following LYAPUNOV function candidate can be found

$$V(x_1, x_2) = x_1^2 + x_2^2. \tag{7.35}$$

Thus the corresponding time derivative is

$$\dot{V}(x_1, x_2) = 2x_1\dot{x}_1 + 2x_2\dot{x}_2 = -2(x_1^2 + x_2^2)^2. \tag{7.36}$$

Hence $V(x_1, x_2)$ is positive definite and $\dot{V}(x_1, x_2)$ is negative definite. Thus the equilibrium at the origin is globally, asymptotically stable for the system.

A quite difficult aspect when using the LYAPUNOV´s direct method is given by how to find LYAPUNOV function candidates. No straight algorithm with a determined solution exists, which is a big disadvantage of this method. SLOTINE [234] proposes several structured approaches to gain LYAPUNOV function candidates namely

- KRASOVSKII´s method,
- the variable gradient method.

Besides these SLOTINE provides additional possibilities to involve the system´s physical principles in the procedure for the determining of LYAPUNOV function candidates while analyzing more complex nonlinear dynamic systems.

7.3.3 Passivity in Dynamic Systems

As another method for the stability analysis of dynamic systems the passivity formalism is introduced within this subsection. Functions can be extended to system combinations by using LYAPUNOV´s direct method, and evaluating the dissipation of energy in dynamic systems. The passivity formalism also is based on nonlinear positive definite storage functions $V(\mathbf{x})$ with $V(\mathbf{0}) = 0$ representing the overall system energy. The time derivative of this energy determines the system´s passivity. As example the general formulation of a system

$$\dot{\mathrm{x}} = \mathrm{f}(\mathrm{x}, \mathrm{u}, t)$$
$$\mathrm{y} = \mathrm{g}(\mathrm{x}, \mathrm{u}, t).$$

is considered. This system is passive concerning the external supply rate $S = \mathbf{y}^T \mathbf{u}$ if the inequality condition

$$\dot{V}(\mathbf{x}) \leq \mathbf{y}^T \mathbf{u} \tag{7.37}$$

is satisfied. KHALIL distinguishes several cases of system passivity depending on certain system characteristics (*Lossless, Input Strictly Passive, Output Strictly Passive, State Strictly Passive, Strictly Passive*) [135]. If a system is passive concerning the *external supply rate S*, it is stable in the sense of LYAPUNOV.

The combination of passive systems using parallel or feedback structures inherits the passivity from its passive subsystems. With the close relation of system passivity to stability in the sense of LYAPUNOV, the examination of the system stability is possible by verifying the subsystem's passivity. Based on this evaluation it can be concluded that the overall system is passive. Always with the assumption that a correct system structure was built.

As an illustrating example the RLC circuit taken from [135] is anaylzed in the following. The circuit structure is shown in figure 7.17:

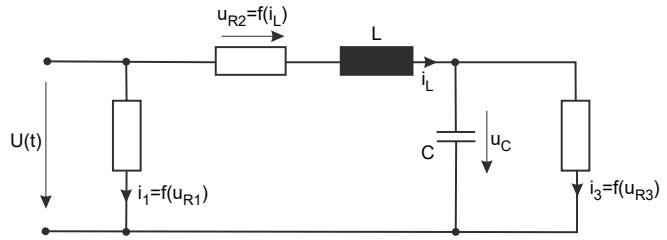


Fig. 7.17 Passivity analysis of an RLC-Network.

The system´s state vector is defined by

$$i_L = x_1$$
$$u_C = x_2.$$

The input u represents the supply voltage U, as output y the current i is observed. The resistors are described by the corresponding voltage current characteristics:

$$i_1 = f_1(u_{R1})$$
$$i_3 = f_3(u_{R3})$$

For the resistor which is coupled in series with the inductor the following behavior is assumed

$$U_{R2} = f_2(i_L) = f_2(x_1). \tag{7.38}$$

Thus the nonlinear system is described by the differential equation:

$$L\dot{x_1} = u - f_2(x_1) - x_2$$
$$C\dot{x_2} = x_1 - f_3(x_2)$$
$$y = x_1 + f_1(u)$$

The presented RLC circuit is passive as long as the condition

$$V(\mathbf{x}(t)) - V(\mathbf{x}(0)) \leq \int_0^t u(\tau)y(\tau)d\tau \tag{7.39}$$

is satisfied. In this example the energy stored in the system is described by the storage function

$$V(\mathbf{x}(t)) = \frac{1}{2}Lx_1^2 + \frac{1}{2}Cx_2^2. \tag{7.40}$$

Equation 7.39 leads to the condition for passivity:

$$\dot{V}(\mathbf{x}(t), u(t)) \leq u(t)y(t) \tag{7.41}$$

which means, that the energy supplied to the system must be equal or higher than the time derivative of the energy function. Using $V(\mathbf{x})$ in the condition for passivity provides

$$\begin{aligned}
\dot{V}(\mathbf{x}, u(t)) &= Lx_1\dot{x_1} + Cx_2\dot{x_2} \\
&= x_1\Big(u - f_2(x_1) - x_2\Big) + x_2\Big(x_1 - f_3(x_2)\Big) \\
&= x_1\Big(u - f_2(x_1)\Big) + x_2 f_3(x_2) \\
&= \Big(x_1 + f_1(u)\Big)u - uf_1(u) - x_1 f_2(x_1) - x_2 f_3(x_2) \\
&= uy - uf_1(u) - x_1 f_2(x_1) - x_2 f_3(x_2)
\end{aligned}$$

and finally

$$u(t)y(t) = \dot{V}(\mathbf{x}, u(t)) + uf_1(u) + x_1 f_2(x_1) + x_2 f_3(x_2). \tag{7.42}$$

In case that f_1, f_2 and f_3 are passive subsystems, i.e. all functions describing the corresponding characteristics of the resistors exist only in the first and third quadrant, so $\dot{V}(\mathbf{x}, u(t)) \leq u(t)y(t)$ is true, hence the RLC circuit is passive. Any coupling of this passive system to other passive systems in parallel or feedback structures again results in a passive system. For any passivity analysis and stability evaluation this method implements a structured procedure and shows a very high flexibility.

In conclusion it is necessary to mention, that all methods for stability analysis introduced in this section show certain advantages and disadvantages concerning their applicability, information value and complexity, regardless whether linear or nonlinear systems are considered. When a stability analysis is expected to be done, the applicability of a specific method should be checked individually. This section only can give a short overview on the introduced methods and techniques, and does explicitly not claim to be a detailed description due to the limited scope of this section. For any further study the reader is invited to consult the proposed literature.

7.4 Control Design for Haptic Systems

As introduced in the beginning of this chapter, control design is a fundamental and necessary aspect within the development of haptic systems. Besides the techniques for system description and stability analysis the need for control design and the applicable design rules become obvious. Especially for the control design of a haptic system it is necessary to deal with several aspects and conditions to be satisfied during the design process. The control design of the haptic device has the focus on the analysis of the control loop closed by the user. In a system for telemanipulation this focus must be extended to the manipulated process or object. In latter case the efforts to be taken are doubled and the aspects to be considererd become more complex. Especially the mechanical coupling of human, may it be on the user's site or on the process' site or both, demands the use of control structures to establish the stability and performance of the overall system continuously and sufficiently. The following sections present several control structures and design schemes in order to set up a basic knowledge about the toolbox for analytic control design of haptic systems. This also involves some of the already introduced methods for system formulation and stability analysis, as these form the basis for most control design methods.

7.4.1 Structuring of the Control Design

As introduced in chapter 5 various different structures of haptic systems exist, whose characteristics will be repeated in this subsection:

- Open-loop impedance controlled: The user experiences an impression of force which is directly commanded via an open loop based only on a demand value. In chapter 5 the basic scheme of this structure is shown by figure 5.1.
- Closed-loop impedance controlled: As it can be seen in figure 5.3 , the user also experiences an impression of force which is fed back to a controller. Here a specific control design will be needed.
- Open-loop admittance controlled: In this scheme on comparison to the previous methods the user erxperiences an impression of a defined position. In the open loop arrangement this position again is directly commanded based only on a demand value. Figure 5.5 shows the corresponding structure of this haptic scheme.
- Closed-loop admittance controlled: This last version as it is depicted in figure 5.7 shows its significant difference in the feedback of the force the user applies to the interface. This force is fed back to a demand value. This results in a closed loop arrangement that incorporates the user and his or her transfer characteristics. In difference to the closed loop impedance controlled scheme this structure uses a force as demanded value $\underline{S}_F$ compared with the detected $\underline{S}_S$, but the system output is still a position $\underline{x}_{out}$. This results in the fact that the incorporation of the user into the closed loop behavior is more complex than it is in a closed loop impedance controlled scheme.

All of these methods can be basically structured as it is depicted in figure 5.11. In this structure all necessary control loops of the overall telemanipulation system become evident.

- On the haptic display and interface site a control loop is closed incorporating the user which is valid as long as the user´s reaction is fed back to the central interface module for any further data processing or control.
- On the process site also, a closed loop exists if measurable process signals (reactions, disturbances) are fed back to the central interface module for data processing or control.
- Underneath these top-level control loops various subsystem control loop exists which have a major impact on the overall system too. As an example, each electrical actuator will most likely be embedded in a cascaded control structure with current, speed and position control.

It becomes obvious that the design of a control system for a telemanipulation system with a haptic interface is complex and versatile. Consequently a generally valid procedure for control design cannot be given. The control structures must be designed step by step involving the following controller:

1. design of all controllers for the subsystem actuators,
2. design of top level controller for the haptic display and interface,
3. design of top level controller for the process or object manipulator.

This strict separation proposed above might not be the only way of structuring the overall system. Depending on the application and functionality, the purposes of the different controller and control levels might be in conflict to each other or simply

overlap. Therefore it is recommended to set up the underlying system structure and define all applied control schemes corresponding to their required functionality. As examining example the following control structure shall be looked at: The altitude control of a passenger airplane takes into account several process variables e.g. the actual altitude, actual speed, fuel consumption, amount of freight, weather, air traffic control clearances. On this level the controller generates demand values for the control loops underneath that are responsible for all stability issues and maneuvers of the airplane. These control loops affect the actuation systems that will move the necessary elevator, ailerons, rudders, thrust control just to name few.

While looking at the control of haptic systems, a similar structure can be established. For both the control of the process manipulation and the haptic display or interface the central interface module will have to generate demand values for force or position, that are going to be followed by the controllers underneath. These demand values derive from a calculation predefined by designed control laws. To gain such control laws a variety of methods and techniques for structural design and optimization can be applied depending on certain requirements. The following subsections give an overview of typical requirements to closed control loop behavior followed by examples for control design.

7.4.2 Requirement Definition

Besides the fundamental need for system stability with sufficient stability margins additional requirements can be set up to achieve a certain system behavior in a closed loop scheme such as dynamic or precision. A quantitative representation of these requirements can be made by the achievement of certain characteristics of the closed loop step response.

Figure 7.18 shows the general form of a typical closed loop step response and its main characteristics: As it can be seen the demanded value is reached and the basic control requirement is satisfied. Additional characteristics are discussed and listed as follows:

$x_{d,max}$ - maximum overshoot
T_{max} - point of time for $x_{d_{max}}$
T_ε - time frame in which the residuum to the demanded value remains within a predefined scope ε
T_{an} - point of time when the demanded value is reached for the first time

For all mentioned characteristics a quantitative definition of certain requirements is possible. For example the number and amplitude of overshoots shall not extend a defined limit or have a certain frequency spectrum that is of special interest for the control design in haptic systems. As it is analyzed in sections 4.2 and 4.3, the user´s impedance shows a significant frequency range which must not be excited within the control loop of the haptic device. Nevertheless a certain cut-off frequency has to be reached to establish a good performance of the dynamic behavior. All these issues

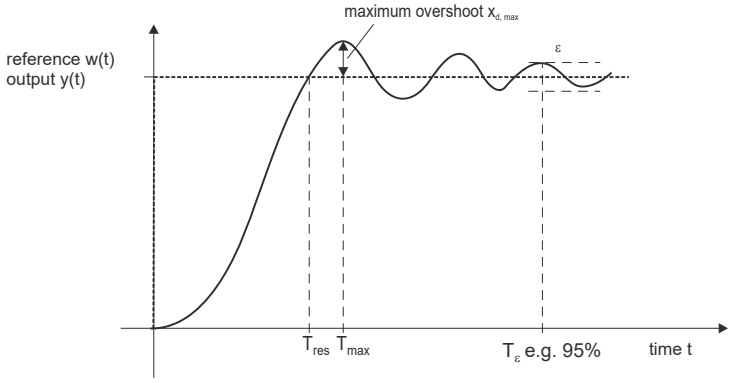


Fig. 7.18 Closed loop step response requirements .

are valid for the requirements to the control design of the process manipulation. In addition to the requirements from the step response due to changes of the setpoint, it is necessary to formulate requirements concerning the closed loop system behavior considering disturbances originating from the process. Especially when interpreting the user´s reaction as disturbance within the overall system description a requirement set up for the disturbance reaction of the control loop has to be established. As it can be seen in figure 7.19 similar characteristics exist to determine the disturbance reaction quantitatively and qualitatively. In most cases both the step response behavior and the disturbance reaction cannot satisfy all requirements, as they often come into conflict with each other which is caused by the limited flexibility of the applied optimization method. Thus it is recommended to estimate the relevance of step response and disturbance reaction in order to choose an optimization approach that is most beneficial. Although determined quantitatively, it is not possible to use all requirements in a predefined optimization method. In most cases an adjustment of requirements is necessary to be made, to apply specific control design and optimization methods. As an example the time T_{An} as depicted above cannot be used directly, and must be transferred into a requirement for the closed loop dynamic characterized by a definite pole placement.

Furthermore simulation techniques and tests offer iteration within the design procedure to gain an optimal control law. However, this very sufficient way of analyzing system behavior and test designed control laws suggests to forget about the analytic system and control design strategy and switch to a *trial an error algorithm*.

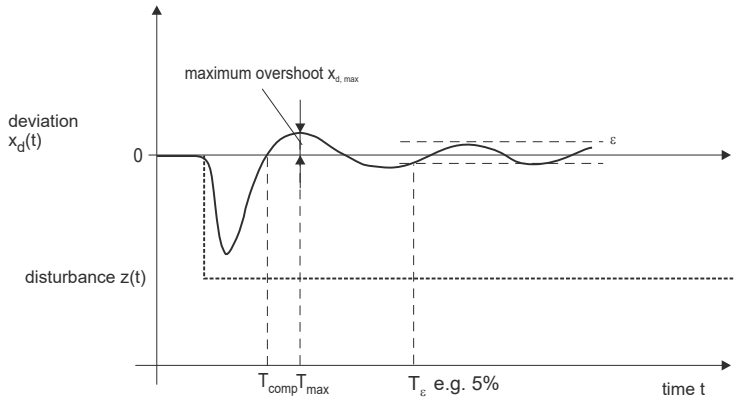


Fig. 7.19 Closed loop disturbance response requirements.

7.4.3 Control Law Design

This section shall present some possible types of controllers and control structures that might be used in the already discussed control schemes. For optimization of the control parameters several methods exists. They are introduced here. Depending on the underlying system description several approaches to set up controllers and control structures are possible. This section will present the classic PID-control, additional control structures e.g. compensation, state feedback controllers, and observer based state space control.

7.4.3.1 Classic PID-Control

Maybe one of the most frequently used controllers is the parallel combination of a proportional (P), an integrating (I) and a derivative (D) controller. This combination is used in several variants including a P-controller, a PI combination, a PD combination or the complete PID structure. Using the PID structure all advantages of the individual components are combined. The corresponding controller transfer function is described by

$$\underline{G_R} = K_R\left(1 + \frac{1}{T_N s} + T_V s\right). \tag{7.43}$$

Figure 7.20 shows the equivalent block diagram of a PID controller structure. Adjustable parameters in this controller are the proportional gain K_R, the integrator time constant T_N and the derivative time T_D.

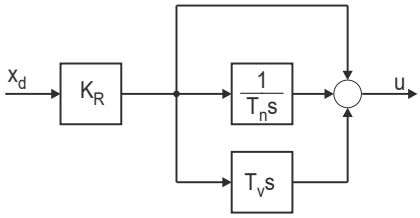


Fig. 7.20 PID block diagram.

With optimized parameter adjustment a wide variety of control tasks can be handled. This configuration offers on the one hand the high dynamic of the proportional controller and on the other, the integrating component guarantees a high precision step response with a residuum $x_d = 0$ for $t \to \infty$. The derivative finally provides an additional degree of freedom that can be used for a certain pole placement of the closed loop system.

As major design techniques the following examples shall be introduced:

- Root locus method: This method has its strength by the determined pole placement for the closed loop system, directly taken into account the dependence on the proportional gain K_R. By a reasonable choice of T_N and T_D the additional system zeros are influenced which affects directly the resulting shape of the root locus and thus the stability behavior. Besides this the overall system dynamic can be designed.
- The second method for the optimization of the closed loop system step response or disturbance reaction by the minimization of an integral criterion is possible. Minimizing the underlying integral criterion the controller parameters are adjusted to optimized values. The basic procedure for this method is as the following: The tracking error x_d due to changes of the demanded set point or a process disturbance is integrated (eventually weighted over time). This time integral will have to be minimized in dependence on all adjustable parameters. In case of convergence of this minimization, the result is the set of optimized controller parameters.

For any additional theoretical background concerning controller optimization the reader is invited to consult the literature on control theory and control design [165], [164].

7.4.3.2 Additional Control Structures

In addition to the described PID controller additional control structures extend the influence on the control result without having an impact on the system stability.

The following paragraphs therefore shall present a disturbance compensation and a direct feedforward of auxiliary process variables.

Disturbance Compensation

The basic principle of disturbance compensation assumes that if a disturbance on the process is measurable and its influence is known, this knowledge can be used to establish compensation by corresponding evaluation and processing. Figure 7.21 shows a simplified scheme of this additional control structure.

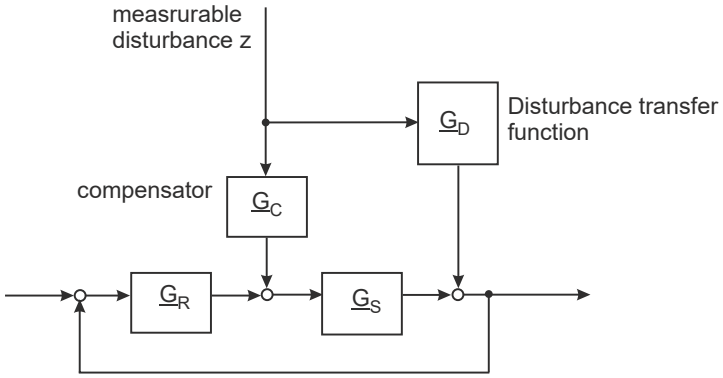


Fig. 7.21 Simplified disturbance compensation.

In this scheme a disturbance signal is assumed to affect the closed loop via a disturbance z transfer function $\underline{G}_{SZ}$. Measuring the disturbance signal and processing the compensator transfer function $\underline{G}_C$ results in a compensation of the disturbance interference. Assuming an optimal design of the compensator transfer function this interference caused by the disturbance is completely erased. The optimal design of a corresponding compensator transfer function is depicted by

$$\underline{G}_C = \frac{-\underline{G}_{SZ}}{\underline{G}_S}. \tag{7.44}$$

This method assumes that a mathematical and practicable inversion of $\underline{G}_{SZ}$ exists. For those cases where this assumption is not valid, the optimal compensator $\underline{G}_C$ must be approximated. Furthermore Figure 7.21 states clearly, that this additional control structure does not have any influence on the closed loop system stability and can be designed independently. Besides the practicability the additional effort should be taken into account. This effort will definitively increase just by the sensors

to measure the disturbance signals and by the additional costs for realization of the compensator.

Auxiliary Input Feedforward

A similar structure compared to the disturbance compensation is the *feedforward* of auxiliary input variables. This principle is based on the knowledge of additional process variables that are used to influence the closed loop system behavior without affecting the system stability. Figure 7.22 shows an example of the feedforward of the demanded setpoint w to the controller signal u using a *feedforward filter function* $\underline{G}_{FF}$.

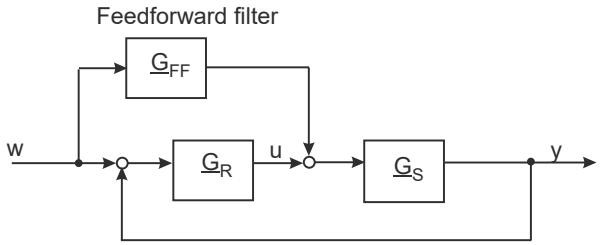


Fig. 7.22 Feedforward of auxiliary input variables.

7.4.3.3 State Space Control

Corresponding to the techniques for the description of multi-input multi-output systems discussed prior in this chapter, the state space control provides additional features to cover the special characteristics within those systems. As described in section 7.2.2 multi-input multi-output systems are preferably depicted as state space models. Using this mathematical formulation enables the developer to implement a control structure that controls the internal system states to demanded values. A big advantage is that the design methods for state space control use an overall approach for control design and optimization instead of a control design step by step for each system state. With this approach it becomes possible to deal with profoundly coupled multi-input multi-output systems with high complexity, and design a state space controller simultaneously. This section will present the fundamental state space control structures. This will cover the *state feedback control* as well as

the *observer based state space control*. For further detailed procedures as well as design and optimization methods the reader is referred to [164] and [265].

State Feedback Control

As it is shown in figure 7.23 this basic structure for state space control uses a feedback of the system states **x**.

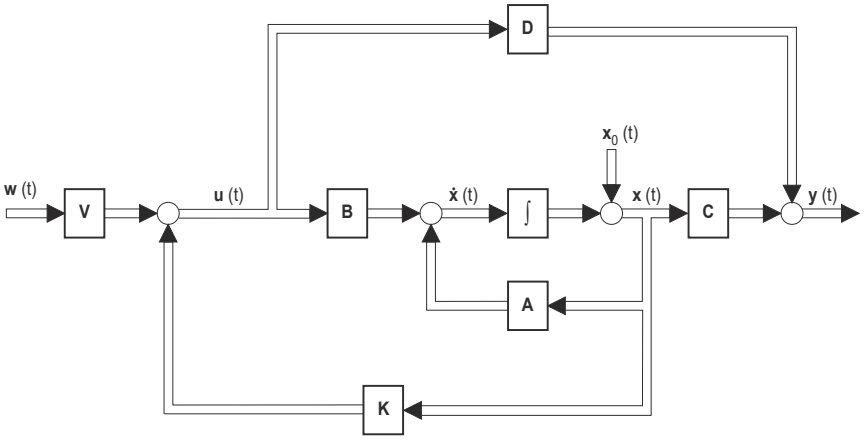


Fig. 7.23 State feedback control.

Similar to the depiction in figure 7.8 the considered system is presented in state space description using the matrices **A**, **B**, **C** and **D**. The system states **x** are fed back gained by the matrix **K** to the vector of the demanded values that were filtered by matrix **V**. The results represent the system input vector u. Both matrices **V** and **K** do not have to be square matrices for a state space description is allowed to implement various dimensions for the state vector, the vector for the demanded values and the system input vector.

Observer Based State Space Control

The state space control structure discussed above requires a complete knowledge of all system states, which is nothing else but that they have to be measured and processed to be used in the control algorithm. From a practical point of view this is not possible all the time due to technical limits as well as costs and effort. As a result the developer is faced with the challenge to establish a state space control without the complete knowledge of the system states. As a solution those system

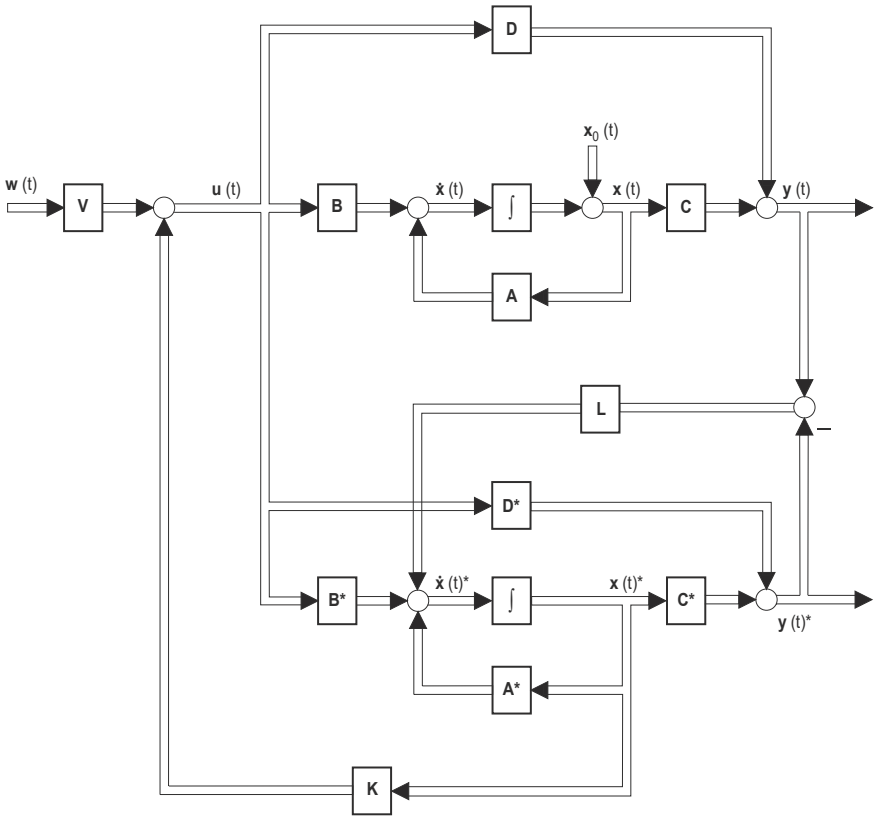


Fig. 7.24 Observer based state space control.

states that cannot be measured due to technical difficulties or significant cost factors are estimated using a state space observer structure that is shown in figure 7.24.

In this structure a system model is calculated in parallel to the real system. As exact as possible this system model is described by the corresponding parameter matrices $\mathbf{A}^*$, $\mathbf{B}^*$, $\mathbf{C}^*$ and $\mathbf{D}^*$. The model input also is represented by the input vector $\mathbf{u}$. Thus the model provides an estimation of the real system states $\mathbf{x}^*$ and an estimated system output vector $\mathbf{y}^*$. By comparison of this estimated output vector $\mathbf{y}^*$ with the real output $\mathbf{y}$, which is assumed to be measurable, the estimation error is fed back gained by the matrix $\mathbf{L}$. This results in a correction of the system state estimation $\mathbf{x}^*$. Any estimation error in the system states or the output vector due to varying initial states is corrected and the estimated states $\mathbf{x}^*$ are used to be gained by the equivalent matrix $\mathbf{K}$ and fed back for control.

This structure of an observer based state space control uses the LUENBERGER observer. In this configuration all real systems states are assumed not to be measur-

able thus the state space control refers to estimated values completely. Practically, the feedback of measurable system states is combined with the observer based estimation of additional system states. In [164] und [265] examples for observer based state space control structures as well as methods for observer design are discussed in more detail.

7.5 Conclusion

The control design for haptic devices faces the developing engineer with a complex manifold challenge. According to the fundamental requirement, to establish a safe reliable and determined influence on all structures, subsystems, or processes the haptic system is composed of, an analytical approach for control system design is not negligible anymore. It provides a wide variety of methods and techniques to be able to cover many issues that arise during this design process. This chapter intends to introduce the fundamental theoretical background. It shows several tasks, functions and aspects the developer will have to focus on, as well as certain methods and techniques that are going to be useful tools for the system's analysis and the process of control design.

Starting with an abstracted view on the overall system, the control design process is based on an investigation and mathematical formulation of the system's behavior. To achieve this a wide variety of methods exists, that can be used for system description depending on the degree of complexity. Besides methods for the description of linear or linearized systems, this chapter introduced techniques for system description to represent nonlinear system behavior. Furthermore the analysis of multi-input multi-output systems is based on the state space description, which is presented here too. All of these techniques on the one hand are aimed at the mathematical representation of the analyzed systems as exact as possible, on the other hand they need to satisfy the requirement for a system description that further control design procedures are applicable to. These two requirements will lead to a trade off between establishing an exact system formulation that can be used in analysis and control design procedures without extending the necessary effort unreasonably.

Within system analysis the overall system stability is the most important aspect that has to be guaranteed and proven to be robust against model uncertainties. The compendium of methods for stability analysis contains techniques that are applicable to linear or nonlinear system behavior, corresponding to their underlying principles that of course limit the usability. The more complex the mathematical formulation of the system becomes, the higher the effort gets for system analysis. This comes in direct conflict to the fact that a stability analysis of a system with a simplified system description can only provide a proof of stability for this simplified model of the real system. Therefore the impact of all simplifying assumptions must be evaluated to guarantee the robustness of the system stability.

The actual objective within establishing a control scheme for haptic systems is the final design of controller and control structures that have to be implemented in the system in various levels to perform various functions. Besides the design of applicable controllers or control structures the optimization of adjustable parameters is also part of this design process As shown in many examples in the literature on control design a comprehensive collection of control design techniques and optimization methods exists, that enable the developer to cover the emerging challenges, and satisfy various requirements within the development of haptic systems as far as automatic control is concerned.

"*This page left intentionally blank.*"

Chapter 8
Kinematic Design

ANDREAS RÖSE

This chapter addresses the basics of the kinematic design of haptic devices. Only few haptic devices generate forces in just one dimension. In such a case the kinematic structure of the device is simple. It is much more challenging and interesting to measure displacements and angles and generate forces and torques in two, three or more degrees of freedom (DOF). For the transformation of forces in numerous spatial directions mechanisms are necessary, which are able to transmit forces from the actuators to the point of action, e.g. the hand. In theory and with some examples two basically different kinematic structures are discussed in this chapter: serial kinematic and parallel kinematic ones.
The chapter focusses on the kinematic design process necessary for haptic applications. It does not cover a universal approach for any form of mechanism. Especially advanced considerations of static and dynamic analysis should be taken from literature recommended within this chapter. Although it excludes advanced optimisation techniques it is still an introduction to the choice and design of kinematic structures for haptic systems.

8.1 Basics

Before taking a look at specific versions of kinematic designs some general terminology and basic mathematical tools have to be understood. This section introduces this knowledge to the reader, starting with simple examples and extending them to a formal analysis of kinematic structures.

T.A. Kern (ed.), *Engineering Haptic Devices*, DOI 10.1007/978-3-540-88248-0_8,
© Springer-Verlag Berlin Heidelberg 2009

8.1.1 Mechanisms and their Classification

Kinematics - the tenet of movement - deals with the mathematical calculation of movement patterns in mechanical systems. In contrast to kinetics, inner and outer forces of the system are neglected. It is a pure concentration on the movement itself. The results of a kinematic observation are ideal transfer characteristics of the observed, massless system.
In general a kinematic structure capable to generate or transfer spatial movements is required. Thus this section deals with the design and kinematic description of - nonlinear transferring - mechanisms with multiple degrees of freedom.
Usually the design process starts with the definition of a topology for the mechanism. Afterwards its transfer characteristics are calculated and - if required - additionally analyzed and optimized. The term “topology” refers to the placement of joints and rigid bodies within the mechanism. Afterwards the geometrical dimensioning follows with respect to specific requirements such as reaching a certain workspace.
A driven mechanism - a robot or robot arm - is made of driven joints - motors, actuators, drives - and un-driven joints - passive joints -. A driven mechanism contains as many actuators, as it has degrees of freedom to move. Special realisations include more actuators than DOFs, what increases the structure’s stiffness. Such designs are not typical for haptic devices and will not be part of this chapter’s focus. Two basic mechanism configurations are distinguished: parallel kinematic and serial-kinematic. A combination of both configurations is called hybrid-kinematic. Figure 8.1 gives examples for all three mechanisms. Purely serial mechanisms include no passive joints; all actuators are in serial order within one single kinematic chain. Parallel mechanisms give the possibility to place all actuators at the frame, minimizing the moving masses. The small inertia makes them highly relevant for haptic applications. The “frame” in the context of kinematic design is a rigid, non-moving body.

Examples

There is a large number of different robots, positioning systems and mechanisms. They all have in common that they are composed of drives and rigid bodies. In some robots the driving power is transmitted via tendons, such as in some haptic devices (fig. 8.2). As the process of dimensioning is identical to all mechanisms two examples are given here, although they do not have an explicit haptic background.
Figure 8.3 shows a fully parallel mechanism with six legs called hexapod by the company Physik Instrumente (PI) GmbH & Co. KG. Within literature this structure is often referred to as STEWARD-GOUGH-platform, as it has been invented by D. STEWARD and V.E. GOUGH and also by K.L. CAPPEL independently from each other at almost the same time. The tool-center-point (TCP) can be moved in six de-

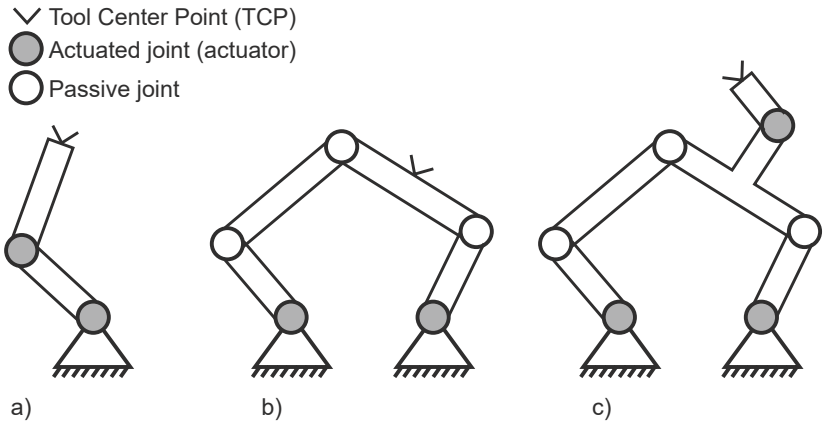


Fig. 8.1 Serial a), parallel b) und hybrid c) mechanism. In the case of a parallel kinematic mechanism the actuators can be placed fixed to the frame.

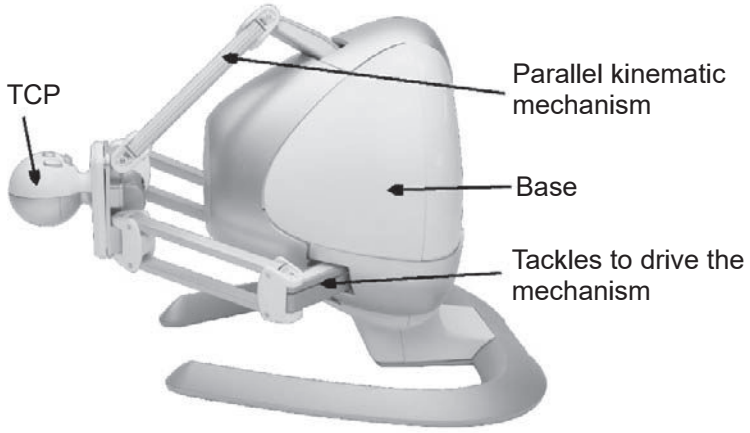


Fig. 8.2 Haptic control element Falcon by the company Novint with parallel kinematic structure. The drives are coupled to the mechanism via tendons. source: Novint Technologies, Inc.

grees of freedom. The movement is generated by six linear actuators connecting the frame to the TCP. Such a structure is easily applicable to haptic devices, as its kinematic relations have been solved analytically [103]. The presented hexapod by the company PI however is a positioning system without the dynamics being necessary for haptic applications.

Figure 8.4 shows a SCARA-robot (Selective Compliance Assembly Robot Arm) by the company Mitsubishi Electric. This kind of robot is used frequently within assembly automation for pick and place tasks. As it is purely serial each actuator has to carry and move all following actuator's weight.

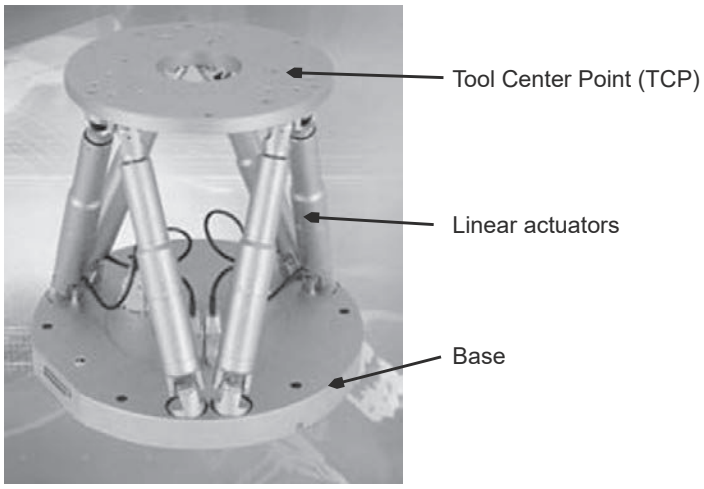


Fig. 8.3 Parallel kinematic hexapod by Physik Instrumente (PI) GmbH & Co. KG. The mechanism contains six DOF, which are driven by liear actuators. source: company PI (www.pi.ws)

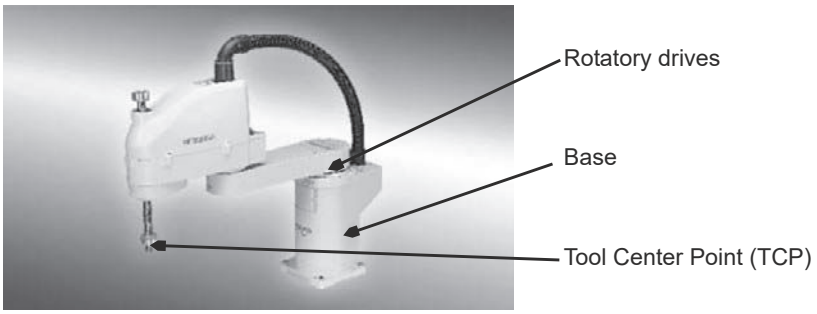


Fig. 8.4 SCARA robot by the company Mitsubishi Electric for use as a pick and place robot. source: Mitsubishi Electric

8.1.2 Calculating Kinematics

The kinematic description of a mechanism for haptic applications can be reduced to the description of in- and output values like *force*, *torques*, *translations* and *angles* at the handle (tool-center-point, TCP) and the actuators.

May $\mathbf{a} = (a_1, a_2, ..., a_n)$ be the vector combining the actuator displacements and angles a_i, and $\mathbf{x} = (x_1, x_2, ..., x_m)$ the vector combining the displacements and angles of the handle x_j. The direct kinematic problem (also: direct kinematics or forward kinematics) can be formulated as

$$\mathbf{x} = f(\mathbf{a}) \tag{8.1}$$

and its inverse - the inverse kinematic problem (also: reverse kinematics or inverse kinematics) - to

$$\mathbf{a} = f^{-1}(\mathbf{x}) \tag{8.2}$$

The forward kinematics according to equation 8.1 is necessary for most haptic applications. It contains the relation for the identification of movements. The hand moves the TCP (**x**-vector) directly. The actuators are moved at the same time via the transmission of the mechanism. For input purposes the position of the TCP shall be identified. If the position of the actuators is measured via sensors (e.g. rotary encoders) the forward kinematics can be used to calculate the position of the hand. Besides position measurement the force output is also based on the forward kinematics too. The calculation of forces can be done by the aid of the Jacobian-matrix (chapter 8.1.3).
For serial mechanisms the derivation of the forward kinematics according to equation 8.1 can quite easily be achieved with the method of DENAVIT-HARTENBERG-parameters (section 8.2.2). For parallel mechanisms the inverse kinematics can frequently be solved by vector analysis, the forward kinematics however is usually challenging to be solved analytically. Luckily for a number of parallel kinematic mehanisms the forward kinematics is already solved within common literature [257].
Annotation:
The before mentioned is relevant for the largest class of haptic devices, which are open- or closed-loop impedance controlled. For open- or closed loop admittance controlled devices the inverse kinematic problem plays an important role. In these cases a change in force is generated by a position change of the TCP relative to the hand. The calculation of in actor positions to achieve a desired change in the position is given by the inverse kinematics.

Example

The Example in figure 8.5 is a Delta-robot. It has a parallel kinematic structure (section 8.3). The Delta Robot has a total number of three DOFs, being able to move in three independent Cartesian coordinates (movements along orthogonal axis). The kinematic problem of Delta-robots is solved and available within literature [257]. It is actuated by three rotatory actuators. When using a Delta-robot for haptic applica-

tions the TCP can be used similar to a joystick for the control of a telemanipulator or simulation. Being used as an impedance controlled parallel kinematic device the forward kinematic solution (equ. 8.1) is necessary to do both, the calculation of the position of the Hand based on angular measurement at the actuators, and actuating the drives to achieve the desired force at the TCP.

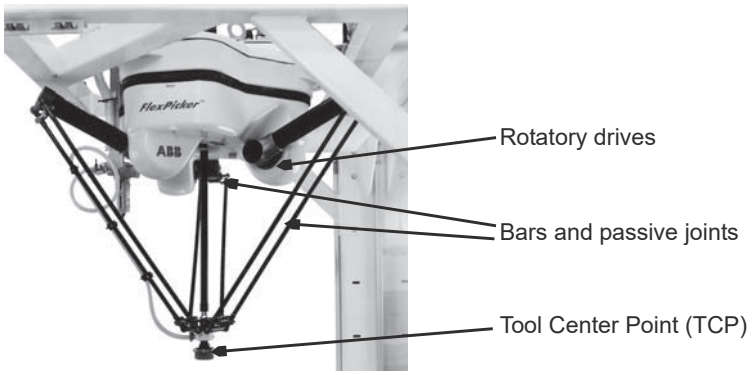


Fig. 8.5 The Delta-Robot has 3 rotational actuators and 3 DOF at its TCP which are best described in catesian coordinates. For a haptic device the TCP would serve as handle. (source: company ABB)

8.1.3 Transfer Characteristics and Jacobian Matrix

To understand this section it is necessary to be aware of the fact that we are dealing with nonlinear transmitting gears with multiple degrees of freedom. Their transfer characteristics - the kinematic dependencies - are impossible to be solved by a system of linear equations. The transmission of displacements and forces changes depending on the position in the workspace. Of course for a haptic device this nonlinear dependency has to be considered and implemented within the control.
When moving a mechanism just a little around an operating point the transfer characteristics can be linearized. The gear in figure 8.6a has a nonlinear transmission function. The input is a rotational drive with angle α and the the output is the displacement x of the driving rod is the output. The forward kinematics is calculated by

$$x = a \cdot sin(\alpha) \tag{8.3}$$

with its functional plot is given in figure 8.6b.

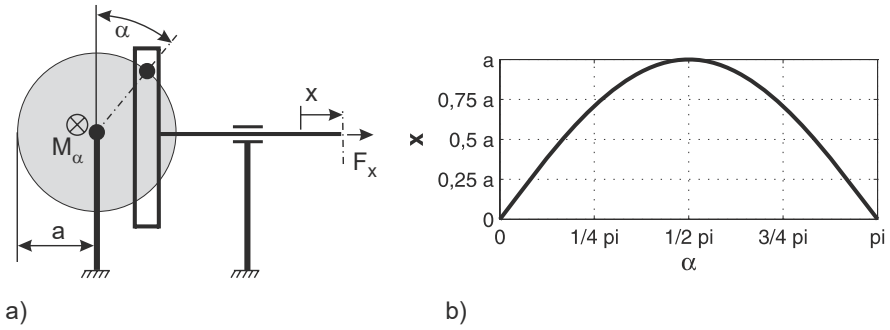


Fig. 8.6 Nonlinear transmitting gear with one degree of freedom. a) functional sceme, b) forward kinematics

Linearizing the function at an arbitrary point α_0

$$dx|_{\alpha_0} = \frac{dx}{d\alpha}|_{\alpha_0} \cdot d\alpha = cos(\alpha_0) \cdot d\alpha = b_0 \cdot d\alpha \tag{8.4}$$

gives the transmission ratio b_0 for differential movements.
The derivation b_0 at α_0 contains the displacement transmission ratio at the input value α_0 as well as the transmission ratio of the forces. Using the principle of virtual work δWx at the drive and δW_α at the transmission bar is calculated:

$$\begin{aligned} \delta W_x &= \delta W_\alpha \\ F_\alpha \cdot \delta\alpha &= F_x \cdot \delta x \\ \frac{F_\alpha}{F_x} &= \frac{\delta x}{\delta\alpha} \\ \frac{M_\alpha \cdot \alpha}{F_x} &= \frac{\delta x}{\delta\alpha} \end{aligned} \tag{8.5}$$

Thus the force-torque transmission ratio at the position α_0 is given by

$$\frac{M_\alpha \cdot \alpha}{F_x} = \frac{1}{b_0} \tag{8.6}$$

which is the reciprocal value of the displacement transfer characteristics. M_α equals the input torque in axial direction of the actuator and F_x the output force in direction x.
Remark:
Virtual values are labelled with a δ, to distinguish them from differential values (d). The principle of virtual work consideres the mechanism in one single position. An infinitesimal small displacement (just some virtual movement) is assumed. By this approach the work is independent from the actual displacement and can be expressed as a force transmission. But even in a differential analysis an infinitesimal

small displacement happens. Consequently the derivation of the displacement is calculated. A singularity is reached when $b_0 = 0$ (e.g. when $\alpha = 90°$). In this position the gear cannot generate defined forces in x-direction anymore which makes it useless for a haptic applications in this point.

Dealing with Multidimensional Problems

In general problems with multiple degrees of freedom have to be analyzed ($\mathbb{R}^n \rightarrow \mathbb{R}^m$). Identical to the one-dimensional situation in these applications a linearization is performed too. The linearization is expressed by the Jacobian matrix $\mathbf{J}$. Considering n degrees of freedom at the actuators $\mathbf{a} = (a_1, a_2, ..., a_n)$ and m degrees of freedom at the tool center points $\mathbf{x} = (x_1, x_2, ..., x_m)$ the Jacobian matrix is calculated by

$$\mathbf{J} = \frac{\partial \mathbf{x}}{\partial \mathbf{a}} = \begin{pmatrix} \partial x_1/\partial a_1 & \cdots & \partial x_1/\partial a_n \\ \vdots & \ddots & \vdots \\ \partial x_m/\partial a_1 & \cdots & \partial x_m/\partial a_n \end{pmatrix} \tag{8.7}$$

The matrix contains all transfer information of the forward kinematics. The inverse Jacobian matrix $\mathbf{J}^{-1}$ includes the same information for the inverse kinematics. The differential displacement behaviour can be expressed by the aid of the Jacobian matrixes as

$$\begin{aligned} d\mathbf{x} &= \mathbf{J} \cdot d\mathbf{a} \\ d\mathbf{a} &= \mathbf{J}^{-1} \cdot d\mathbf{x} \end{aligned} \tag{8.8}$$

The result is a system of linear equations with coefficient matrix $\mathbf{J}$. When the number of actuators is equal to the number of DOF of the TCP, the system of equations has a square coefficient matrix. From linear algebra it is known that a system of equations with a square coefficient matrix contains an unlimited number of solutions if the determinant of the matrix equals 0. Consequently one actuator DOF respectively one DOF of the TCP can be chosen freely if either one of the conditions

$$det(\mathbf{J}) = 0$$

respectively

$$det(\mathbf{J}^{-1}) = 0 \tag{8.9}$$

is fulfilled. In these cases a defined force transmission is no longer possible. In case of a haptic device this is identical to a malfunction of the unit. The positions in which the determinants of $\mathbf{J}$ and $\mathbf{J}^{-1}$ equals zero are called singular positions. The transfer functions have a singularity in this position (in contrast to all other positions

where the functions are regular, i.e. clearly defined). For the dimensioning of a haptic device the workspace has to be chosen such that all singularities are excluded. This requirement is important for the design of haptic devices, especially for parallel mechanism (section 8.3) what makes their design even more challenging compared to serial mechanisms.

Remark on singularities: The specific situation with a non-square Jacobian matrix is always the case when the TCP is characterized in more coordinates than number of drives exist. If there is a mechanism with four actuators controlling four DOFs of the TCP, it is usually helpful to describe the TCPs movement in six coordinates. These are three Cartesian directions of movement and three Euler-angles. The TCP will then move simultaneously in more than one of these coordinates. However the coordinates will not become independently controllable. In this case we achieve a 6×4 Jacobian matrix. The preconditions for a singularity in these situations are given by the rank of the Jacobian Matrix [30]. Is the rank of either $\mathbf{J}$ or $\mathbf{J}^{-1}$ less than 4, the position is singular.

For the derivation of force transfer characteristics a similar approach as in equation 8.5 can be used for multidimensional positioning systems [257]. The vector $\mathbf{f}_a = (f_{a1}, f_{a2}, \ldots, f_{an})^T$ is the vector of all forces and/or torques f_{ai} on the driven joints. The vector $\mathbf{f}_x = (f_{x1}, f_{x2}, \ldots, f_{xm})^T$ is the vector of all forces and torques f_{xj} acting on the tool-center-point. The virtual displacement of the actuators is labelled $\delta\mathbf{a} = (\delta a_1, \delta a_2, \ldots, \delta a_n)$, those of the TCP are named $\delta\mathbf{x} = (\delta x_1, \delta x_2, \ldots, \delta x_m)$. The indices are identical to those in equation 8.7. Calculating the virtual work

$$\begin{aligned} \delta W_x &= \delta W_a \\ \mathbf{f}_x^T \cdot \delta\mathbf{x} &= \mathbf{f}_a^T \cdot \delta\mathbf{a} \end{aligned}$$

with the relation $d\mathbf{x} = \mathbf{J} \cdot d\mathbf{a}$ from equation 8.8 yields

$$\mathbf{f}_x^T \cdot \mathbf{J} \cdot \delta\mathbf{a} = \mathbf{f}_a^T \cdot \delta\mathbf{a}$$

The concept of virtual work is applicable to any non-singular configuration for any virtual displacement $\delta\mathbf{a}$. Consequently

$$\mathbf{f}_x^T \cdot \mathbf{J} = \mathbf{f}_a^T$$

is valid. By transposing the equation the final solution is

$$\mathbf{f}_a = \mathbf{J}^T \cdot \mathbf{f}_x \tag{8.10}$$

giving the relation between forces and torques on the actuators and the tool-center-point. The same approach generates the inverted equation

$$\mathbf{f}_x = \mathbf{J}^{-T} \cdot \mathbf{f}_a \tag{8.11}$$

Remarks:

- Equation 8.10 is of highest interest for the control of haptic devices. It explicitly contains the input-output-relation of forces. When a specific force/torque condition shall be applied at the TCP this equation calculates the necessary control values for the actuators.
- Equation 8.10 and 8.11 are calculated from a purely kinematic analysis and neglect dynamic forces of inertia or friction and static forces of e.g. gravity. For haptic devices with comparably heavy mechanics these forces have to be considered [257]. This is always the case when the forces from accelerated masses or gravity get into ranges which cannot be neglected anymore. Example: If a mass of $m = 5\,g$ is accelerated linear in $t = 0.1\,s$ over a distance of $s = 1\,mm$ a force of $F = m \cdot s/t^2 = 0.5\,N$ can be reached easily.

8.1.4 Optimizing the Transfer Characteristics

A specific topology consisting of actuators, rods and passive joints within a multidimensional transmitting gear shows a nonlinear transfer characteristic, which is influenced mainly by the aspect ratios of its components. According to MERLET parallel mechanisms with well designed dimensions can perform better than one with better suited topology but worse dimensions [174]. Optimization in this context is defined as the change of specific geometrical dimensions within the same mechanism to achieve a specific performance regarded as optimum. This may be a maximized workspace with homogeneous transfer characteristics of forces from the TCP to the actuators. A systematic and frequently numerical optimization is applied mainly on parallel kinematic mechanisms. In direct comparison serial kinematic displays can be designed according to geometrical requirements quite easily. Additionally singularities of serial kinematic chains are obvious. They appear when the kinematic chains are stretched. Nevertheless the transfer characteristics of serial kinematics can still be altered and optimized by changing the length of rods.
To perform an optimization the following steps should be taken:

1. Definition of parameters and their span of values (e.g. rod lengths)
2. Analytical or numerical description of the optimization problem
3. Mathematical optimization, e.g. via a gradient approach or evolutionary algorithms

These steps are discussed in [185] and covered in detail within [138]. The key challenge is the formal description of the optimum. In [24] several optimizations are given using the singular values of the Jacobian matrix as a measure. Within theses approaches the measurement value for an optimum has to be determined by scanning the complete workspace and optimizing relevant parameters of the mechanism between each scanning process. Usually this approach needs a significant amount

of time. In practical work it is highly recommend to intensively study literature for possible analysis and optimization of the chosen mechanism, as an own optimization will mostly need unreasonable efforts and time. The optimization of complex mechanisms with multidimensional optimization and evolutionary algorithm is still - although often discussed - subject of research work and not necessarily successful in every case.

8.2 Serial Mechanisms

This section describes the approach for a design of a serial kinematic structure for haptic devices. The design is structured in three steps:

- Topological synthesis - the design of an assembly of rods and actuators
- Calculation of the kinematics - Mathematical description of displacements provided by the user and the output of forces
- Dimensioning - Defining the geometrical dimensions (section 8.1.4)

8.2.1 Topological Synthesis

A serial mechanism is not less or more than a sequence of rods and actuators, whereas the actuators can be regarded as driven joints. Whether the actuators are linear or rotary is of no importance for the complexity of the kinematic problem. For the workspace and the orientation of the tool-center-point however this aspect is of highest importance. A spacial serial mechanism with three rotatory drives changes the orientation of its TCP all over its workspace. If it is not intended to generate a torque as output to the user, the handle attached to this serial mechanism has to be equipped with a passive universal joint. Such a realisation as haptic device can be found in figure 8.7. Torques are decoupled from the hand. The handle does not have to be placed exactly in the TCP, as the moments are eliminated by the passive joints. Force vectors can be displaced arbitrarily within space. As a result the Hand experiences the same forces like the TCP.

When dimensioning a serial haptic device several aspects should be considered during the selection of joints:

- Linear joints tend to cant with loads orthogonal to the direction of movement. This requires either a design minimizing orthogonal forces on linear joints, or these joints must be designed adequately to work under orthogonal load.
  In an idle situation and in the case of impedance-controlled devices actuated joints are moved passively by the user. This subordinates them to the same requirements as passive linear joints.

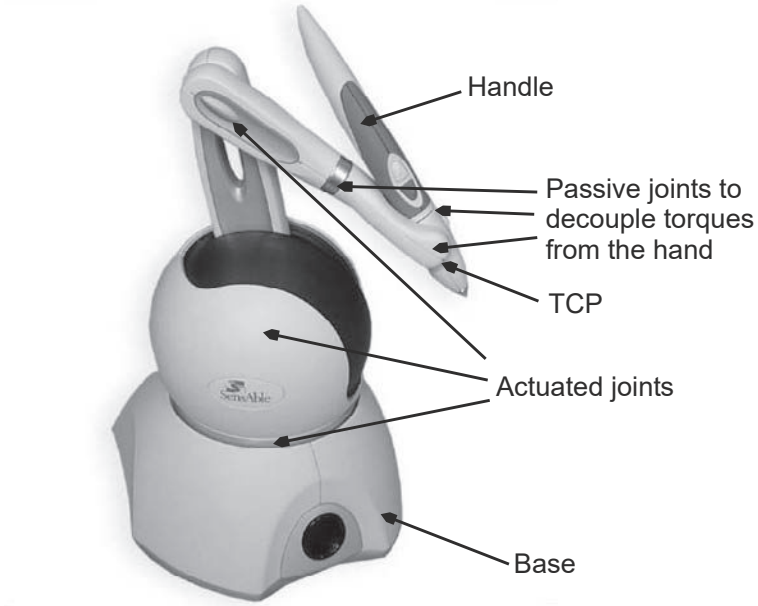


Fig. 8.7 The PHANTOM Omni haptic device by SensAble Technologies is an example for a spacial working serial kinematic haptic device. The hand is decoupled from rotational movements by passive joints. Thus no torques are induced to the hand.

- Rotatory joints are not able to transmit forces, when their kinematic chain is in a stretched position. Within the example in figure 8.6 this situation happens for $\alpha = \pi/2$. The passive joint at the outer area of the disk reaches a stretched position. As obvious from equation 8.5 no defined haptic feedback can be generated in this position. The same applies to the forces in radial direction to the prismatic joint.
- A general recommendation for impedance-controlled devices is the minimization of friction within the joints. One option to achieve this is the use of ball bearings for rotational joints. When used for linear joints the roll motion is unsteady and therefore noticeable.

As human beings are equipped with many serial kinematic chains (e.g. arms, legs) the working area of a serial kinematic chain can be understood intuitively. This makes it simple to design a corresponding haptic control unit. This is however not the only criterion and will be further addressed in section 8.4.
The design can be done geometrically "with circle and ruler", however the following should be considered:

- Actuators add inertia and moving masses to the mechanism. In serial mechanisms all actuators are placed in series. This has negative influence on the dynamics of force transmission. Approaches for a dynamic analysis can be found in [257] and [102].

- A simple design criterion could be, to place actuators as near as possible to the base platform of the system and use transmission elements to the points/joints of actuation.

8.2.2 Calculation of the Kinematic Problems

Section 8.1.3 introduced the importance of kinematic transfer functions for the design of haptic devices. Calculating the force transfer characteristics the forward kinematics is necessary for impedance controlled devices, and the inverse kinematics for admittance controlled devices. The following section describes a method to calculate the forward kinematics for any serial mechanism called the "DENAVIT-HARTENBERG-parameters" method. Literarily this method is based on climbing along a serial kinematic chain step by step. This is done by multiplying a homogeneous coordinate transformation matrix containing the coordinates of each element with the combined product of its predecessors. This transformation matrix is dominated by four parameters, called the DENAVIT-HARTENBERG-parameters. Mathematically moving through the whole kinematic chain and using this approach the kinematic correlation between the TCP and the frame - the forward kinematics - is obtained.
An adequate homogeneous coordinate transform matrix is given by the matrix **T** [102]:

$$\mathbf{T} = \begin{pmatrix} \mathbf{R} & \mathbf{t} \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} R_{1,1} & R_{1,2} & R_{1,3} & t_1 \\ R_{2,1} & R_{2,2} & R_{2,3} & t_2 \\ R_{3,1} & R_{3,2} & R_{3,3} & t_3 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{8.12}$$

with homogeneous coordinate transformation matrices rotations, translations, scaling and perspective transformations can easily be calculated [257]. For robotic applications only rotations and translations are relevant, which are included in the specific matrix given in equation 8.12.
Remarks:

- Matrix multiplications are associative, but they are not commutative. Thus the order for multiplication is highly important for calculating the forward kinematics (see below).
- The numbers (0, 1) in the last row of the matrix guarantee that rotations and translations do not influence each other. With this feature a simple algorithm can perform rotations and translations with a single matrix multiplication. This increases the clarity of an implementation and may be one reason why homogenous coordinate transformations are widespread within robotics and even virtual reality programming.

Using the homogenous transformation matrix $\mathbf{T}$ a vector $\mathbf{p}$ - itself being homogenous too - can be transformed via scalar multiplication from one coordinate system (0) in another coordinate system (1):

$$\mathbf{p}_1 = \mathbf{T} \cdot \mathbf{p}_0 \tag{8.13}$$

In order to perform multiplications with a 4×4 matrix the vectors have to be of type 4×1, which is mandatory for homogenous vectors. For robotic applications a vector $\mathbf{p}$ is formed as

$$\mathbf{p} = (p_x, p_y, p_z, 1)^T \tag{8.14}$$

where p_x, p_y and p_z are Cartesian coordinates.
The calculation of the forward kinematics happens in the following steps:

1. Definition of a coordinate system for each body attached to its lower end beginning at the coordinate system attached to the frame and ending in a coordinate system attached to the TCP. The coordinate systems have to be oriented in such way that the coordinate system (i+1) can be reached by its predecessor coordinate system (i) with the combination of the following changes:
   - Rotation θ around the z_i-axes (rotation according to the right-hand rules, clockwise looking in the direction of the rotational axes)
   - Displacement d along the z_i-axes
   - Displacement a along the x_{i+1}-axes
   - Rotation α along the x_{i+1}-axes (right-hand rule)
2. Composition of a table with all DENAVIT-HARTENBERG-parameters (θ, d, a, α) of the whole chain. The chain includes n coordinate systems.
3. Formulation of a transformation matrix ($\mathbf{T}_i$) for each group of DENAVIT-HARTENBERG-parameters (see below).
4. Multiplication of the transformation matrixes to calculate the total transformation $\mathbf{T}_{tot} = \mathbf{T}_n \cdot \ldots \cdot \mathbf{T}_i \cdot \ldots \cdot \mathbf{T}_0$

The resulting total transformation matrix $\mathbf{T}_{tot}$ transforms the base coordinate system $\mathbf{p}_0$ to the TCP coordinate system $\mathbf{p}_{TCP}$ and represents the forward kinematics.

$$\mathbf{p}_{TCP} = \mathbf{T}_{tot} \cdot \mathbf{p}_0 \tag{8.15}$$

Each matrix $\mathbf{T}_i$ is composed of the DENAVIT-HARTENBERG-parameters by the following scheme:

$$\mathbf{T}_i = \begin{pmatrix} cos(\theta_i) & -sin(\theta_i) \cdot cos(\alpha_i) & sin(\theta_i) \cdot sin(\alpha_i) & a \cdot cos(\theta_i) \\ sin(\theta_i) & cos(\theta_i) \cdot cos(\alpha_i) & -cos(\theta_i) \cdot sin(\alpha_i) & a \cdot sin(\theta_i) \\ 0 & sin(\alpha_i) & cos(\alpha_i) & d \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{8.16}$$

The matrix $\mathbf{T}_{tot} = \prod_{i=n}^{0} \mathbf{T}_i$ depends on every Denavit-Hartenberg-parameter. It is the serial execution of $n+1$ transformations in three dimensional space with six degrees of freedom. Therefore $\mathbf{T}_{tot}$ is given by

$$\mathbf{T}_{tot} = \begin{pmatrix} cos\alpha \cdot cos\beta & cos\alpha \cdot sin\beta \cdot sin\gamma - sin\alpha \cdot cos\gamma & cos\alpha \cdot sin\beta \cdot cos\gamma + sin\alpha \cdot sin\gamma & x \\ sin\alpha \cdot cos\beta & sin\alpha \cdot sin\beta \cdot sin\gamma + cos\alpha \cdot cos\gamma & sin\alpha \cdot sin\beta \cdot cos\gamma - cos\alpha \cdot sin\gamma & y \\ -sin\beta & cos\beta \cdot sin\gamma & cos\beta \cdot cos\gamma & z \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{8.17}$$

Where x,y and z are the Cartesian coordinates of the TCP relative to the origin. α, β und γ are roll-, pitch- and yaw angle, equivalent to a rotation around the x-axis (roll, α) a rotation around the new y-axis (pitch, β) and a rotation around the resulting z-axis (yaw, γ). They can be extracted out of the matrix by some simple calculations. Instead of roll-, pitch- and yaw angle other conventions like Euler-angles can be used, of course. Accordingly the interpretation of the matrix $\mathbf{T}_{tot}$ will have to change.
In contrast to the direct kinematics the inverse kinematics is hard to derive for serial kinematic problems. Some methods can be obtained from [257]. The inverse kinematics is only relevant for admittance based devices and of minor importance for impedance-based devices.

8.2.3 Example of a Serial Mechanism

Figure 8.8 gives an example for a serial mechanism with four driven joints q_1 to q_4, three of them rotary and one linear. Coordinate systems are located in the centres

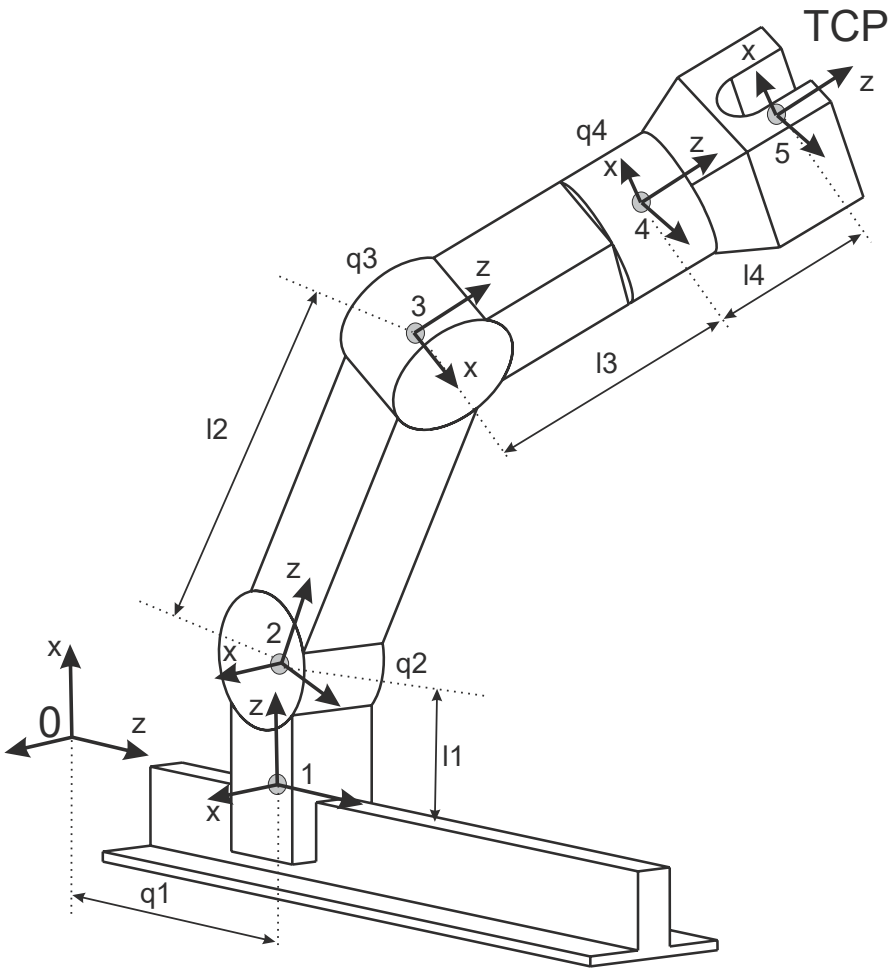


Fig. 8.8 Example of a serial mechanism

of each actuator such as in the TCP (1 to 5). As a reference a world coordinate system (0) marks a fixed point on the frame. The coordinate systems are placed such that each of them can be reached by a combined movement according to the four DENAVIT-HARTENBERG-parameters from the preceding one (section 8.2.2). In this example the forward kinematics of the mechanism shall be calculated. First the necessary parameters are collected within table 8.1.
then the transformation matrixes T_0 to T_4 can be generated.

Table 8.1 DENAVIT-HARTENBERG-parameters of the mechanism in fig. 8.8

n	θ (rotation around z_n)	d (displacement in z_n-direction)	a (displacement in x_{n+1}-direction)	α (rotation around x_{n+1})
0	90°	q_1	0	90°
1	0	l_1	0	$-q_2$
2	90°	l_2	0	$-q3$
3	$-q_4$	l_3	0	0
4	0	l_4	0	0

$$\mathbf{T}_0 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & q_1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\mathbf{T}_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & cos(-q_2) & -sin(-q_2) & 0 \\ 0 & sin(-q_2) & cos(-q_2) & l_1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\mathbf{T}_2 = \begin{pmatrix} 0 & -cos(-q_3) & sin(-q_3) & 0 \\ 1 & 0 & 0 & 0 \\ 0 & sin(-q_3) & cos(-q_3) & l_2 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\mathbf{T}_3 = \begin{pmatrix} cos(-q_4) & -sin(-q_4) & 0 & 0 \\ sin(-q_4) & cos(-q_4) & 0 & 0 \\ 0 & 0 & 1 & l_3 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\mathbf{T}_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & l_4 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

The total forward-kinematics T_{tot} is a result of the product

$$T_{tot} = T_4 \cdot T_3 \cdot T_2 \cdot T_1 \cdot T_0$$

and is dependent only on the geometrical dimensions (l_1, l_2, l_3, l_4) and the positions of the driven joints (q_1, q_2, q_3, q_4).
Remarks:

- The resulting analytical solution will become quite bulky due to the large numbers of multiplications of sine and cosine terms, but will be fast to computate with given values for q_1 to q_4.

- The Jacobian-matrix of the forward kinematics will be even more complex, but will have to be calculated for the force transfer characteristics of impedance based devices. A recommendable help for the calculation are programs for the solution of symbol based mathematical problems like the symbolic math toolbox of MatLab®.

8.3 Parallel Mechanisms

In contrast to serial assemblies parallel kinematic mechanisms consist of coupled passive and active joints. In general each active joint (each actuator) contributes to one degree of freedom of the TCP's movement. Hence all previous considerations from section 8.1.3 can be applied to parallel mechanisms. This section coveres the design of parallel kinematic mechanisms which is devided in the following steps:

- Topological synthesis - the design of an assembly of odys and joints
- Calculation of the kinematics - Mathematical description of displacements provided by and the output of forces to the user
- Dimensioning - Defining the geometrical dimensions (section 8.1.4)

8.3.1 Topological Synthesis

The GRÜBLER Formula

If active and passive joints are coupled by bodies with the aim, that the resulting mechanism offers as many degrees of freedom as it contains actuators, then the number of degrees of freedoms from the joints will have to be exactly defined and distributed according to specific rules. The absolute number of passive joint DOFs to be distributed can be calculated by the formula of GRÜBLER which can easily be described [174][102][24]:
A mechanism with n rigid bodies (including the frame) without a coupling of the elements in a three-dimensional space has a total number of

$$F = 6 \cdot (n-1)$$

degrees of freedom. Each body has six DOFs in space. The frame itself is fixed to space, it has no DOF. When coupling any free body with g joints of DOF number f_i ($i = 1 \ldots g$) F is reduced by $(6 - f_i)$. Thus F can be calculated as

$$F = 6 \cdot (n-1) - \sum_{i=1}^{g} (6 - f_i)$$
$$= 6 \cdot (n-g-1) + \sum_{i=1}^{g} f_i \tag{8.18}$$

and is called gear-DOF. F is the number of relative movements which have to be constrained, to constrain the whole mechanism. Calculating the number of joint-DOFs according to equation 8.18 a mechanism results, which can be actuated by F actuators in F DOFs. Within the calculation the so called identical conditions f_{id} and constraints s have to be corrected to exactly define the number of joints.

$$F = 6 \cdot (n-g-1) + \sum_{i=1}^{g} f_i - f_{id} + s \tag{8.19}$$

F	Gear-DOF
n	Number of bodys
g	Number of joints
f_i	Degree of freedom of the ith joint
f_{id}	Sum of identical condition
s	Sum of constraints

Remark:
An identical condition is given for example when a rod has universal joints at both of its ends. The rod will be able to rotate around its axis, without violationg any constrains. Another example are two coaxial oriented linear joints. Constraints appear whenever conditions have to be fulfilled to enable the movement [121]. If 5 joint axes have to be parallel to a 6th axis to enable a movement, then $z = 5$. Another example for a passive condition are two driving rods which have to be placed in parallel to enable a motion.

Calculation of the Joint DOFs to be distributed

Usually in the state of the design of the mechanism topology the number of gear-elements and joints is not yet defined. The origin for the design process is the question about the total degrees-of-freedom F the mechanism has to achieve. This value for F is identical to the total number of actuated joints. A parallel kinematic mechanism can be classified according to its level of parallelism [102]. If the mechanism contains as many legs as it contains DOFs - an example is the hexapod in figure 8.3 - then it is called fully parallel. Other designs are called partly parallel or high-grade parallel. In this context the focus lies on fully parallel designs. The correlation between the number of chains k, joints g and elements n is defined according to

$$n = g - k + 2 \tag{8.20}$$

Frame and TCP are taken into account by the right hand summand “2”. By insertion into equation 8.19 a relation between chain number k and gear-DOF F and the sum of all joint-DOFs $f_{tot} = \sum_{i=1}^{g} f_i$ is achieved.

$$f_{tot} = \sum_{i=1}^{g} f_i = F + 6 \cdot (k-1) + f_{id} - s \tag{8.21}$$

Remarks:

- For fully parallel mechanisms every chain - a connection between the frame to the TCP - contains exactly one actuator with one DOF.
- The number of chains k is identical to the number of DOFs for fully parallel mechanisms. Hence equation 8.21 simplifies to $f_{tot} = 7 \cdot F - 6 + f_{id} - s$ in this case

Kinematic Scheme and Joint Distribution

A “kinematic scheme” is used to clarify the distribution of joint-DOFs (including the actuators) within the design of a mechanism. The kinematic scheme of a Delta-robot (fig. 8.5) is shown in figure 8.9. This robot consists of $k = 3$ chains, has a total number of $F = 3$ DOFs and does not include identical DOFs nor any constraints. According to equation 8.21

$$f_{tot} = F + 6 \cdot (k-1) = 15$$

joint DOFs have to be distributed among three chains. This is often done symmetrically - such as in this case too. Each chain consists of 5 DOFs. Each leg includes one rotary actuator representing one DOF. The remaining DOFs are realized as universal joints resulting in a joint-distribution as shown in figure 8.9b.

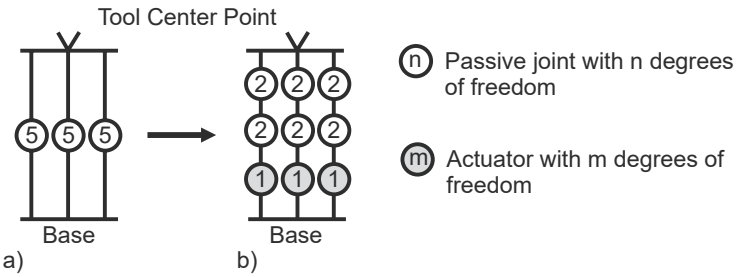


Fig. 8.9 Kinematic scheme of the delta robot from fig. 8.5. The mechanism contains 15 joint-DOF distributed to 3 legs (a). Every leg contains a rotary actuator with 1 DOF (b). The joint DOFs are denoted in the circles.

Many other distributions of joint-DOFs are possible and will result in a different

topology of the mechanism, of course. Usually the design process itself is quite intuitive. As a simple rule of thumb the following advices can be used:

- Concentrate joint-DOFs: Universal joints and ball joints simplify the design and the transfer characteristics compared to serially placed joints with just one DOF each. Especially the kinematic transfer characteristics are much easier to formulate with concentrated joints. Additionally the play within joints will be concentrated which will reduce the play of the total mechanism.
- Use actuators which themselves are fixed to the frame: The drives of the hexapod are placed in the rods and move during positioning of the platform. Due to the little workspace covered this does not influence the dynamics of the hexapod significantly. However still with the actuators being fixed to the frame the dynamic properties of the design would be increased, as their mass does not count to the inertia of the device anymore.
- DOFs should be distributed symmetrically: A symmetrical design should be preferred compared to an asymmetrical one. The kinematic transfer functions get significantly simpler. For an asymmetrical design it is very likely that no closed analytical solution will be found.

For the first draft of a parallel kinematic haptic system it is recommend basing on known parallel kinematic mechanisms. Generally there is a wide range of known systems in literature. Examples of typical structures can be found in [185] [24] and [174].

8.3.2 Calculation of the Kinematic Transfer Functions

In contrast to serial kinematics the inverse kinematic problem of a parallel mechanism is easier to calculate than the forward kinematics. Even with simple designs it is challenging or even impossible to analytically solve the direct kinematics, which is the base for the force transfer function of impedance based haptic devices. Generally speaking systems with symmetries and concentrated degrees of freedom within passive joints are easier to solve than other designs. In [259] and [258] analytical solutions of the forward kinematic problem of simple manipulators with three degrees of freedom are shown. Other designs can be found in [257] and [103]. The procedure of calculating the inverse kinematic problem can be split up into the following three steps:

1. Formulation of closed vector chains for each leg, starting at the reference coordinate system (0) enclosing the TCP coordinate system (0) and going back to the reference coordinate system..
2. Splitting the vector chains in all - Cartesian - movement directions of the individual leg.
3. Solving the resulting system of equations according to the TCP's coordinates.

An example for this calculation is given in section 8.3.3. The transfer function of parallel kinematic mechanisms and especially their forward kinematic calculation include singularities (section 8.1.3). Usually they are a result of stretched orientations of passive joints. A problem that is obvious within serial mechanisms. It is mandatory to consider this problem during device design.

8.3.3 Examples of a Parallel Mechanism

Figure 8.10 shows one example of a planar parallel mechanism. The actuated joints are located at q_1 and q_2 and summarized to the vector **q**. The lengths of both rods on the right side are identical (a) and are placed at the same distance on both sides (b), resulting in a parallel linkage. This prevents the TCP from rotating and makes it move within planar Cartesian coordinates $\mathbf{x} = (x, y)$. The corresponding coordinate system is labelled with T.

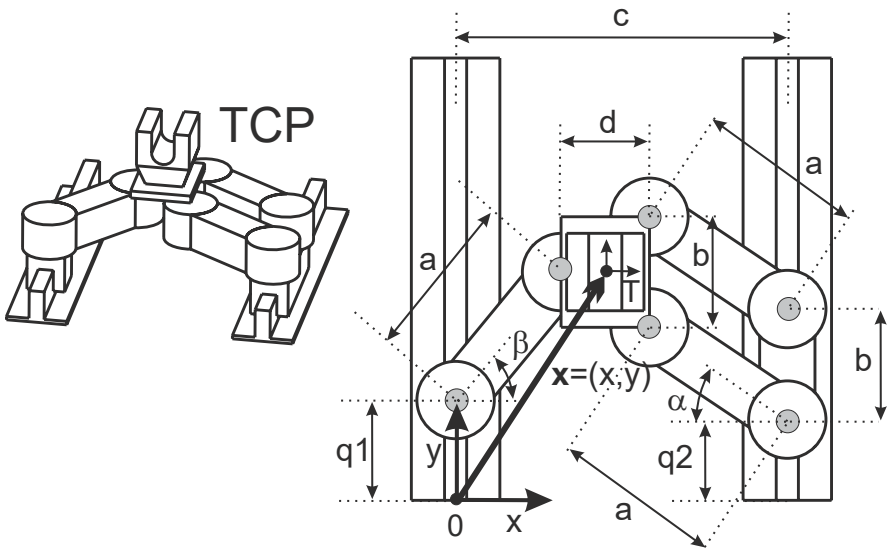


Fig. 8.10 Example of a parallel mechanism

Following the GRÜBLER method for the total number of mechanism degrees of freedom (equ. 8.21) this parallel linkage generates a constraint. Additional five constraints are given by the requirement, that all five axis of the rotary joints are placed in parallel to the axis of the 6th rotatory joint. Resulting in $s = 6$ constraints. The mechanism is composed of $k = 3$ chains, is expected to have $F = 2$ degrees of freedom (movement in the x-y-plane without rotating). Identical DOFs (section 8.3.1) are not existent ($f_{id} = 0$). Equation 8.21 calculates

$$f_{tot} = F + 6 \cdot (k-1) + f_{id} - s = 8$$

DOFs, which have to be distributed among the joints to achieve a mechanically defined mechanism. Two of these DOFs are given by the linear drives q_1 and q_2. The remaining 6 DOFs have to be realized as passive joints. Here, this is achieved by 6 passive rotational joints. The following paragraph deals with the calculation of the inverse kinematic problem $\mathbf{q} = f^{-1}(\mathbf{x})$, which is not too complicated in this case. The inverse kinematics would be suitable for an admittance controlled haptic device. It can be calculated with the aid of two closed vector chains in world-coordinates via the tool-center-point. These vector chains include both legs one the right and left side and are split up according to their x- and y-coordinate.

$$x = c - a \cdot cos(\alpha) - \frac{d}{2} \Rightarrow sin(\alpha) = \sqrt{1 - \left(\frac{c - \frac{d}{2} - x}{a}\right)^2} \tag{8.22}$$

$$y = q_2 + a \cdot sin\alpha + \frac{b}{2} \Rightarrow q_2 = y - a \cdot sin(\alpha) - \frac{b}{2} \tag{8.23}$$

$$x = a \cdot cos(\beta) - \frac{d}{2} \Rightarrow sin(\beta) = \sqrt{1 - \left(\frac{x - \frac{d}{2}}{a}\right)^2} \tag{8.24}$$

$$y = q_1 + a \cdot sin\beta \Rightarrow q_1 = y - a \cdot sin(\beta) \tag{8.25}$$

By combining 8.22 with 8.23 and 8.24 with 8.25 the solutions for q_1 and q_2 can be found:

$$q_1 = y - a \cdot \sqrt{1 - \left(\frac{x - \frac{d}{2}}{a}\right)^2}$$

$$q_2 = y - a \cdot \sqrt{1 - \left(\frac{c - \frac{d}{2} - x}{a}\right)^2} - \frac{b}{2}$$

These are the inverse kinematic transfer functions.
Remarks:

- Additionally the solution of the forward kinematic problem $\mathbf{x} = f(\mathbf{q})$ is necessary for calculating the position of the TCP based on position measurements taken at the actuators. If the forward kinematic problem is not known, one option could be to include additional sensors within the haptic device to gain additional information about position and orientation.
- The solution of the equations include $sin(\phi) = \sqrt{1 - cos^2(\phi)}$. The radicand of the square root has to stay positive for real solutions. This is a plausible indicator that the mechanism contains singularities.

8.4 The Complete Process of Kinematic Design

Within this chapter the basics for the kinematic design process were discussed. This final section will summarize the whole process and repeat it in context. Figure 8.11 shows the rough structure.

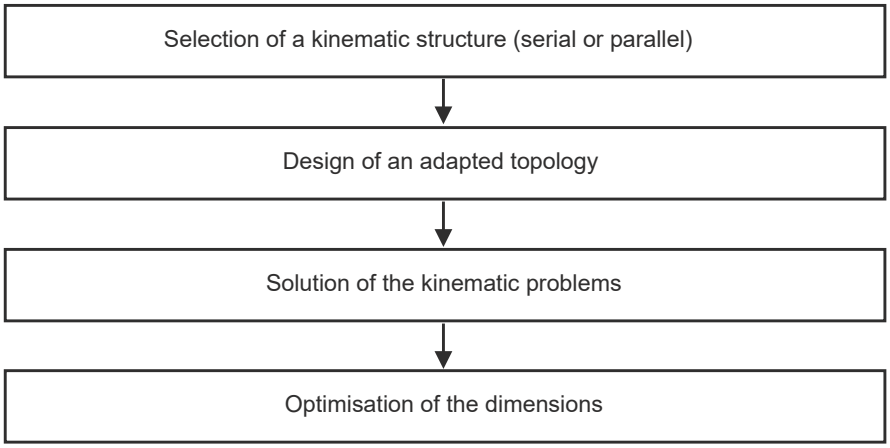


Fig. 8.11 The complete process of kinematic design.

Selection of a Structure

Basically two different structures exist: “serial kinematic” or “parallel kinematic”. A mix of both is called “hybrid kinematic”, combining some degrees of freedom with parallel elements, while others are purely serial. This variant combines the advantages and disadvantages of both concepts. The haptically relevant properties of each variant are given in table 8.2 and compared qualitatively. Serial and parallel kinematics were discussed in section 8.2 and 8.3.

Table 8.2 Interesting properties of serial kinematic and parallel kinematic mechanisms for haptic applications

Property	Serial Mechanisms	Parallel Mechanisms
Mechanical impedance	high	low
Ratio workspace/size	high	low
Calculation of the forward kinematic problem	easy	difficult

The most important criterion of table 8.2 is given with the mechanical impedance. For haptic devices - especially open- or closed-loop impedance controlled devices - the mechanism's impedance should be as low as possible. This guarantees a good transformation of driving forces to the TCP. The easiest way to achieve low impedance is given with parallel kinematics with light rods and joints with low friction. The actuators should be fixed to the frame in a parallel concept and located near the frame in a serial mechanism. A transmission of movement energy via tendons should be considered. The haptic device "Delta.3" from the company ForceDimension (fig. 8.12) with the kinematic structure of a delta-robot (fig. 8.5) fulfils these recommendations and is therefore an example of a well chosen kinematic structure. In the haptic device "Falcon" of the company Novint (fig. 8.2) the same kinematic configuration is chosen with variants in the design of joints and the quality of components. The intensive usage of the delta structure shows its suitability for haptic applications. Concerning the concepts of open- or closed-loop control (section 7) all kinematic principles are suitable to be used.



Fig. 8.12 The haptic device "Delta.3" by the company ForceDimension uses a parallel kinematic structure with rotatory actuators fixed to the frame. The bodys are constructed lightweight to achieve a low mechanical impedance.

Design of a Topology

The design process of an adequate topology has been discussed within the section 8.2.1 and 8.3.1. Frequently the topology can be defined intuitively when dealing with serial structures.
The topology-design for parallel structures is more challenging. First the number of necessary DOFs have to be identified and a kinematic scheme has to be designed or selected (section 8.3.1). Afterwards a single leg stretching from the frame to the tool center point should be dimensioned. Adding additional legs to the first one con-

strains movement capabilities of the mechanism, for example by constrained parallel movements. For fast progress especially with parallel mechanism a detailed review of existing literature of industrial robots is recommended [174] [185] [257].

Calculation of the kinematic problems

Depending on the design of the topology the direct kinematics may be analytically unsolvable. Even the inverse kinematics, the kinematic description for admittance based devices, may not be analytically solvable in some cases. Section 8.2.2 gives a general method for the calculation of forward kinematics of serial mechanisms. When an analytical solution can not be found, numerical real-time solutions are an interesting option. A simple Netwon-Raphson-approximation method is able to find roots for closed vector chains. But even on a PC with several GHz of computing power these approximation algorithms are not fast enough for haptic applications yet. Adjusted hardware, e.g. field programmable gate arrays (FPGA) are suitable for these types of calculations and are able to solve such kinematic problems at an update rate considerably below 1 ms.

Optimization of Structural Dimensions

Section 8.1.4 gave already a short insight into the challenges for automated optimization of mechanisms. Even if no complete optimization can be done, a fundamental thing to do is an analysis for singularities within the workspace, and an intuitive and iterative optimization of the structure to guarantee a force transmission in any point of the workspace.

Chapter 9
Actuator Design

THORSTEN A. KERN, MARC MATYSEK, STEPHANIE SINDLINGER

The preceding chapters were focused on the basics of control engineering and kinematic design. They covered topics of structuring and fundamental character. This and the following chapters deal with the design of technical components as parts of haptic devices. Actuators are the most important elements of every haptic device, as their selection respectively their design influences the quality of the haptic impression significantly. This chapter deals with frequently used actuators structured according to their physical working principle. Each actuator type is discussed as to its most important physical basics, with examples of their dimensioning, and one or more applications given. Other, rarely used actuator principles for haptic systems are mentioned in section 9.6 “special forms”, or are classified in the overview of physical working principles in section 9.1.

Experience teaches us that actuators for haptic applications can rarely be found “off-the-shelf”. Their requirements always include some outstanding features in rotational frequency, power-density or geometry. These specialties make it necessary and helpful for their applicants to be aware of the capabilities and possibilities of modifications of existing actuators. Hence this chapter addresses both groups of readers: the users who want to choose a certain actuator and the mechanical engineer who intends to design a specific actuator for a certain device from scratch.

T.A. Kern (ed.), *Engineering Haptic Devices*, DOI 10.1007/978-3-540-88248-0_9,
© Springer-Verlag Berlin Heidelberg 2009

9.1 General Facts about Actuator Design

THORSTEN A. KERN

Before a final selection of actuators is made, the appropriate kinematics and the control-engineering structure, according to the previous chapters, should have been fixed. However, in order to handle these questions in a reasonable way, some basic understanding about actuators is mandatory. Especially the available energy densities, forces and displacements should be estimated correctly for the intended haptic application. This section provides some suggestions and guidelines to help and preselect appropriate actuators according to typical requirements.

9.1.1 Overview of Actuator Principles

There are a certain number of approaches to transform an arbitrary energy source into mechanical energy. Each of these ways is one actuation principle. The best known and most frequently used principles are:

- **Electrodynamic principle** - A force, so called LORENTZ-force, acting upon a conductor conducting a current
- **Electromagnetic principle** - A force, acting upon a magnetic circuit to minimize the enclosed energy
- **Piezoelectric principle** - A force, acting upon the atomic structure of a crystal and deforming it.
- **Capacitive principle** - A force, resulting from charges trying to minimize the energy in a capacitor
- **Magnetorheological principle** - Viscosity change within a fluid resulting from particles trying to minimize the energy contained within a magnetic circuit.
- **Electrochemical principle** - displacement of or pressure within a closed system, whereby a substance emits or bounds a gas and consequently changes its volume due to the application of electrical energy.
- **Thermal principle** - Change of length of a material due to controlled temperature changes, making use of the material's coefficient of thermal expansion.
- **Shape-memory alloy** - Sudden change of an object's shape made of special material due to relatively small temperature changes ($\approx 500°C$). The object transforms into a root shape embossed during manufacture by the application of high temperatures ($\approx 1000°C$).

Each of these principles is used in different embodiments. They mainly differ in the exact effective direction of e.g. a force vector[1] or a building principle[2]. As a consequence a wide-spread terminology exists for naming actuators. The major terms are given here:

- **Electric motor** - the most general term at all. It may describe any electromechanic transformer. However, in most cases it refers to an actuator rotating continuously whose currents are commutated (mechanically or electronically), or which is equipped with a multiphase alternating current unit. Typically, it is a synchronous motor, a drive with a rotor moving synchronously to the rotating electromagnetic field. In a more general understanding the term includes hysteresis motors and squirrel-cage rotors, too. But the latter has not yet reached any relevance for haptic systems, even not in very exceptional cases.
- **EC-motor** - Specific embodiment of the synchronous motor and very common to haptic applications. Motor according to the electromagnetic or electrodynamic principle with an electronic control unit for the rotating field (electronic-commutated, electronically commutated).
- **DC-motor** - Another specific form of a synchronous motor and used among haptic applications because of its cheapness and simplicity. This is an actuator according to the electromagnetic or electrodynamic principle with a mechanical control unit for the rotating filed using switching contacts (mechanically commutated).
- **Resonance-actuator** - Generic term for a whole class of actuators with very different actuation principles. The term describes an actuator containing one component which is driven in its mechanical resonance mode (or nearby its resonance mode). Typically, parts of this component make an elliptic oscillation driving a second component in small steps via a frictional coupling. As a result of the high frequency, the movement of the second component seems uninterrupted. The term is most frequently applied to piezoelectric actuators.
- **Ultrasonic-actuator** - Resonance-actuator performing steps at a frequency within ultrasonic ranges (>15kHz). These actuators are built almost always according to the piezoelectric principle.
- **Voice-coil-actuator** - Drive according to the electrodynamic principle. Mainly consisting of a conductor wrapped around a cylinder. The cylinder itself is placed in a magnetic circuit, resulting in a Lorentz-force when a current is applied to the conductor. There are two major embodiments, one with a “moving coil”, another variant with a “moving magnet”.
- **Shaker** - Another form of a voice-coil-actuator with an elastic suspension of the coil. When current is applied to the coil an equilibrium condition between the suspension’s spring and the Lorentz-force is achieved at a specific displacement. Actuators according to this structure are frequently used for fast and dy-

[1] The electromagnet principle for instance is divided into magnetic actuators and actuators according to the reluctance principle; the piezoelectric principle is subdivided into three versions depending on the relative orientation of electrical field and movement direction.

[2] e.g. resonance drives vs. direct drives

namic movements of masses for vibration testing (this is where its name comes from). It is nothing else but a loudspeaker principle without a spatially expanded membrane for sound generation. Another embodiment of this principle primarily aiming at measurement applications is called ampere balance. In this case a given load is applied against the elasticity resulting in a displacement. The current is increased until the displacement is reset to zero. The value of the current is a direct measure for the load applied. This measurement principle of forces/torques is also known as "THOMSON balance" or "electrodynamic balance".

- **Plunger-type magnet** - Actuator according to the electromagnetic principle. A rod made of ferromagnetic material is pulled into a magnetic circuit equipped with a coil. These actuators have very nonlinear force-displacement characteristics.
- **Stepper-motor** - Generic term for all actuation principles moving forward step by step. In contrast to the resonance-drives no component of a stepper-motor is driven in resonance mode. Their step-frequency is below any resonance of the system and may vary. These actuators may even be used in a "microstep mode", interpolating movement between so called "full-steps", which are original to their mechanical design. The term is most frequently used for rotatory drives. Especially for those working according to the reluctance-principle or another electromagnetic actuation principle.
- **Pneumatic and Hydraulic** - These actuation principles do not have a direct electric input value. They transform pressure and volume-flow into displacement and force. The media for pressure transmission is air in case of pneumatics and a fluid, typically oil, in case of hydraulics. Usually, the pressure itself is generated via actuators, e.g. electrical actuators attached to a compressor, and controlled via electrical valves.
- **Bending-actuator** - Actuator with an active layer, frequently made of a piezoelectric material - attached to a passive mechanical substrate. By actuating the active layer, mechanical tensions between this layer and the substrate build up, resulting in a bending movement of the whole actuator.
- **Piezoelectric stack** - A larger number of piezoelectric layers mechanically connected in series. Small displacements of each layer sum up to a large usable displacement of the whole actuator.
- **Piezoelectric motor/drive** - Generic term for all drives according to the piezoelectric principle. It frequently refers to drives moving a rotor or translator with a frictional coupling. However, this movement does not need to happen in resonance mode.
- **Capacitive actuator** - Actuator according to the capacitive principle and frequently used in microtechnology/micro engineering. Usually equipped with a comb-like structure of electrode-pairs, generating forces in milli-newton with micrometers of displacement.
- **Shape-memory wire** - Wire on the basis of shape-memory alloys capable to shorten in the range of percents ($\approx 8\%$ of its total length) when changing its temperature (e.g. by controlling a current flowing through the wire. The current heats up the wire according to its thermal loss at the wire's electrical resistance).

- **Surface-wave actuators** - Generic term for a group of actuators generating high-frequency waves in mechanical structures or exciting the resonance-modes of structures. This actuator is frequently based on piezoelectric principles and has been used for the generation of haptic textures for some years.

Each of the above actuation principles can be found in tactile and/or kinaesthetic systems. To simplify the decision process for a new design, all actuators can be grouped into classes. Most of the physical working principles can be grouped either into "self-locking" or "free-wheeling" systems. These groups are identical to:

- Positional-sources (x) respectively angular-sources (α)
- Force-sources (F) respectively torque-sources (M)

According to the basic structures of haptic systems (chapter 5) it is likely that both classes are used within different haptic systems. The correlation between basic structures of haptic systems and actuators is depicted in table 9.1. This table shows a tendency towards typical applications; however by adding mechanical elements (springs, dampers) it is possible to use any actuator for any basic structure of haptic systems.

Table 9.1: "Classic" Application areas for actuator principles in haptic systems

X: is frequently used by many groups and even commercialized;
(X): some designs, especially in research;
-: very rare to almost none, an if it is used, it is only in the context of research);
TYPE: Gives an idea about which actuator design (translatory or rotatory) is used more often. If the actuator is unusual but does exist, the marker is set into brackets.

Control-type:		**Admittance**		**Impedance**	
Type	Actuator	closed-l.	open-l.	open-l.	closed- l.
Rot.	electric motor[3]	X	X	(X)[4]	-
Rot. & transl.	EC-motor	-	-	X	X
Rot & transl.	DC-motor	-	-	X	X
Rot & transl.	Resonance-actuator	X	X	(X)	-
Rot & transl.	Ultrasonic-actuator	X	X	(X)	-
Transl.	Voice-Coil	-	-	X	X
Transl.	Shaker	X	X	-	-
Transl.	Plunger-type magnet	X	-	-	-
Rot. (& transl.)	Stepper-motor	X	X	-	-
Transl. (& rot.)	Pneumatic	(X)	X	-	-
Transl. (& rot.)	Hydraulic	-	X	-	-
Transl.	Bending-actuator	-	X	-	-

[3] in the meaning of a mechanically commutated drive with a power between 10 to 100 W
[4] by high frequency vibrations of the commutation

Control-type:		Admittance		Impedance	
Type	Actuator	closed-l.	open-l.	open-l.	closed- l.
Transl.	Piezo-stack	(X)	X	-	-
Transl. & rot.	Piezo-actuator	X	X	X	-
Transl.	Capacitive	-	(X)	-	-
Transl.	Shape-memory	-	(X)	-	-
Transl.	Surface wave	-	(X)	-	-

9.1.2 Actuator Selection Aid Based on its Dynamics

Different actuator designs according to the same physical principle still cover wide performance ranges regarding their achievable displacements or forces. Based on the author's experience, these properties are put into relation to the dynamical ranges relevant for haptic applications. In figure 9.1 the most important actuation principles are visualized in squares scaled according to their achievable displacements (a)[5] and typical output forces and torques (b). The area covered by a specific version of an actuator is typically smaller than the area shown here. The diagram should be read in such a way that e.g. for haptic applications, . electromagnetic linear actuators exist, providing a displacement up to 5 mm at $\approx$ 50 Hz. These designs are not necessarily the same actuators which are able to provide $\approx$ 200 N, as with electromagnetic systems the effectively available force increases with smaller maximum displacement (section 9.3). The diagrams in figure 9.1 visualize the bandwidths of realization-possibilities according to a certain actuation principle and document the preferred dynamic range of their application. Using the diagrams, we have to keep in mind that the borders are fluent and have to be regarded in the context of the application and the actuator's individual design.

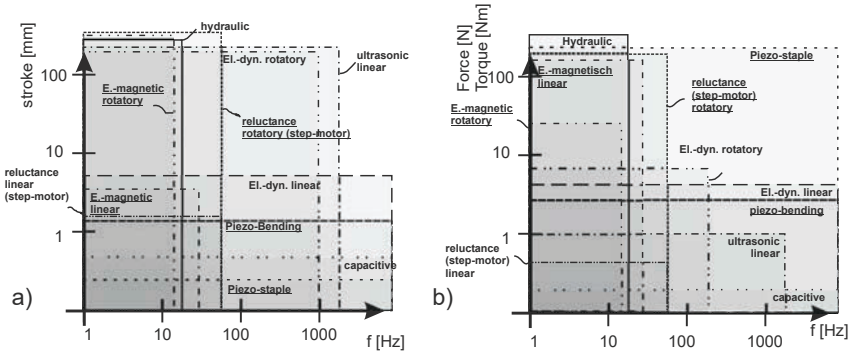


Fig. 9.1 Order of the actuator principles according to the achievable displacement (a) and forces resp. torques (b) in depedency from their dynamics.

[5] for continuously rotating principles all displacements are regarded as unlimited

9.1.3 Gears

In general machine engineering the usage of gears is a matter of choice for adapting actuators to their load and vice versa. Gears are available in many variants. A simple lever can be a gear; complex kinematics according to chapter 8 are a strongly nonlinear gear. For haptic applications specialized gear designs are discussed for specific actuation principles in the corresponding chapters. However, there is one general aspect of the application of gears with relevance to the design of haptic systems which has to be discussed independently: the scaling of impedances.

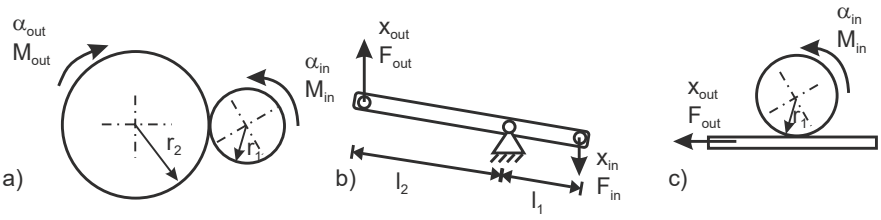


Fig. 9.2 Simple gear design with wheels (a), a level (b) and a Bowden-cable (c).

There is no principal objection to the use of gears for the design of haptic systems. Each gear (fig. 9.2) may it be rotatory/rotatory (gearwheel or frictional wheel), translatory/translatory (lever with small displacements), rotatory/translatory (Bowden-cable) has a transmission "*tr*". This transmission ratio scales forces and torques neglecting loss due to friction according to

$$\frac{F_{\text{out}}}{F_{\text{in}}} = tr = \frac{l_2}{l_1}, \tag{9.1}$$

$$\frac{M_{\text{out}}}{M_{\text{in}}} = tr = \frac{r_2}{r_1}, \tag{9.2}$$

$$\frac{F_{\text{out}}}{M_{\text{in}}} = tr = \frac{1}{2\pi r_1}, \tag{9.3}$$

and displacements resp. angles according to

$$\frac{x_{\text{in}}}{x_{\text{out}}} = tr = \frac{l_2}{l_1}, \tag{9.4}$$

$$\frac{\alpha_{\text{in}}}{\alpha_{\text{out}}} = tr = \frac{r_2}{r_1}, \tag{9.5}$$

$$\frac{\alpha_{\text{in}}}{x_{\text{out}}} = tr = \frac{1}{2\pi r_1}. \tag{9.6}$$

The velocities and angular velocities scale analogously to the differential of above equations. Assuming the impedance of the actuator $\underline{Z}_{\text{transl}} = \frac{\underline{F}}{\underline{v}}$ resp. $\underline{Z}_{\text{rot}} = \frac{\underline{M}}{\underline{\alpha'}}$, one consequence of the load-condition of a driven impedance $\underline{Z}_{\text{out}}$ from the perspective of the motor is:

$$\underline{Z}_{\text{transl}} = \frac{\underline{F}_{\text{in}}}{\underline{v}_{\text{in}}} = \frac{\underline{F}_{\text{out}}}{\underline{v}_{\text{out}}} \frac{1}{tr^2} = \underline{Z}_{\text{transl out}} \frac{1}{tr^2} \tag{9.7}$$

$$\underline{Z}_{\text{rot}} = \frac{\underline{M}_{\text{in}}}{\underline{\alpha'}} = \frac{\underline{M}_{\text{out}}}{\underline{\alpha'}} \frac{1}{tr^2} = \underline{Z}_{\text{rot out}} \frac{1}{tr^2} \tag{9.8}$$

The transmission-ratio tr is quadratic for the calculation of impedances. From the perspective of an actuator, the driven impedance of a system is getting small with a gear with a transmission-ratio larger than one. This is favorable for the actuating system (and the original reason for the application of gears). For haptic applications, especially for impedance controlled ones, the opposite case has to be considered. In an idle situation and with a transmission ratio larger than one[6] the perceived mechanical impedance of a system $\underline{Z}_{\text{out}}$ increases to the power of two with the transmission-ratio. Another aspect makes this fact even more critical, as the increase of output force changes only in a linear way with the transmission ratio, whereas e.g. a motor's unwanted moment of inertia is felt to increase quadratically. This effect is obvious to anyone who has ever tried to rotate a gear-motor with a high transmission ratio (e.g. $tr > 100$) at its output. The inertia and the internal frictions upscaled by the gear is identical to a self-locking of the actuator.

As a consequence, the usage of gears with force-controlled haptic systems makes sense only for transmission ratios of 1 to 20 (with emphasis on the lower transmission ratios between 3 to 6). For higher transmission ratios gear designs according to figure 9.2c and equation 9.6 have proved valid. They are used in many commercial systems, as with the aid of the definition $tr = \frac{1}{2\pi r_1}$ and the included factor 2π a comparably high transmission ratio can be achieved easily. In combination with rotatory actuators (typically EC-drives) with low internal impedances this design shows very impressive dynamic properties. Some advice may be given here out of practical experience: Gearwheels are applicable for haptic systems but tend to generate unsteady and waving output torques due to their toothing. By a careful mechanical design, this unsteadiness can be reduced. The mechanical backlash should be minimized (which is typically accompanied by an increase in friction) for example by material combinations with at least one soft material. At least one gear should have spur/straight gearing, whereas the other one can keep involute gearing.

[6] which is the normal case, as typically the fast movement of an actuator is transmitted into a slower movement

9.2 Electrodynamic Actuators

THORSTEN A. KERN

Electrodynamic actuators are the most frequently used type of drives for haptic applications. This popularity is a result of the direct proportion between their output value (force or torque) from their input values (the electrical current). In case of kinaesthetic applications they are typically used as open-loop controlled force sources. But even with tactile applications these very dynamic actuators are frequently applied for oscillating excitements of skin areas. They are equally often used rotatory and translatory. Depending on the design either the electrical coil or the magnet is the moving component of the actuator. This section gives a short introduction to the mathematical basics of electrodynamic systems. Afterward some design variants are discussed in more details. The final subsection deals with the drive electronics necessary to control electrodynamic systems.

9.2.1 The Electrodynamic Effect and its Influencing Variables

Electrodynamic actuators are based on the LORENTZ-force

$$\mathbf{F}_{\text{Lorentz}} = \mathbf{i} \cdot l \times \mathbf{B}, \tag{9.9}$$

acting upon moving charges in magnetic fields. The LORENTZ-Force is dependent on the current **i**, the magnetic induction **B** such as the length of the conductor l, which is typically formed as a coil. This subsection deals with optimization of each parameter for the maximisation of the generated output force F_{Lorentz}. Any electrodynamic actuator is made of three components:

- generator of the magnetic filed (coil, or most frequently a permanent magnet)
- magnetic flux conductor (iron circuit, magnetic core)
- electrical conductor (frequently formed as coil or a more complex winding)

After a shallow look a recommendation for the maximization of the output force could be to simply increase the current **i** in the conductor. However with a given and limited available space for the conductor's length l (coil's cross section), and a flux density **B** with an upper border (0.8 to 1.4 T), the effectiveness of this change has to be put into question. This can be shown with a simple calculation example:

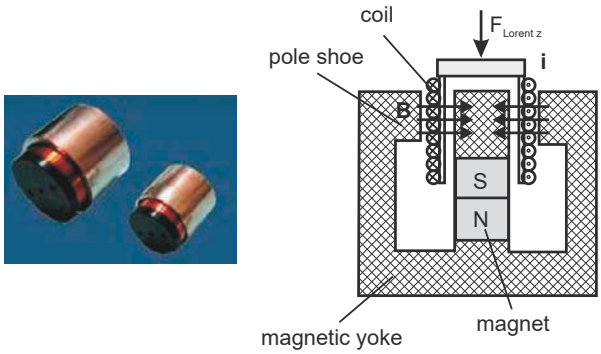


Fig. 9.3 Moving-coil actuator AVM by adrive-technology building the basis for the following example.

9.2.1.1 Efficiency Factor of Electrodynamic Actuators

For example a straight-forward design of an electrodynamic actuator similar to the AVN 20-10 (fig. 9.3) is analyzed. It contains a wound coil with a permanent-magnet in a ferromagnetic core. The electrical power loss P_{el} of this electrodynamic system is generated mainly in a small moving coil with a pure ohmic resistance $R_{\mathrm{coil}} = 3.5\,\Omega$ and a nominal current $i = 0.78\,A$:

$$P_{\mathrm{el}} = R_{\mathrm{coil}}\, i^2 = 3.5\,\Omega \cdot 0.78\,\mathrm{A}^2 = 2.13\,\mathrm{W}. \tag{9.10}$$

With this electrical power loss, and at a flux density **B**=1.2 T, and with an orthogonal conductor orientation, and a conductor length within the air gap l=1.58 m, the actuator itself generates the force

$$F_{\mathrm{Lorentz}} = i\,l\,B = 0.78\,\mathrm{A} \cdot 1.58\,\mathrm{m} \cdot 1.2\,\mathrm{T} = 1.48\,\mathrm{N}. \tag{9.11}$$

Assuming the system being driven in idle mode - working against the coil's own mass of m=8.8 g only - being accelerated from idleness, and performing a displacement of x=10 mm, above electrical power P_{el} is needed for a period of

$$t = \sqrt{2\frac{x}{a}} = \sqrt{2\frac{x\,m}{F}} = 0.011\,s \tag{9.12}$$

seconds. The electrical energy loss sums up to

$$W_{\mathrm{el}} = P_{\mathrm{el}} \cdot t = 23,4\,mJ. \tag{9.13}$$

This gives an efficiency factor of $\frac{W_{\mathrm{mech}}}{W_{\mathrm{el}}+W_{\mathrm{mech}}} = 38\%$ for idle mode and continuous acceleration. Assuming now that the same actuator shall generate a force of 1 N against a finger tip for a period of e.g. two seconds, an electrical power of

$W_{\mathrm{el}} = 2.13\,\mathrm{W} \cdot 2\mathrm{s} = 4.26\,\mathrm{J}$ is needed. This would be identical to an efficiency factor well below 1%. And indeed the efficiency factor of electrodynamic actuators in haptic applications lies in the area of low percentages due to the common requirement to generate quite static forces without much movement. This simple calculation points to one major challenge with electrodynamic actuators: The electrical power lost due to heat transmission extends the mechanically generated power by far. Consequently during the design of electrodynamic actuators attention has to be paid to the optimization of power within the given actuator volume and to the thermal management of energy losses.

9.2.1.2 Minimization of the Power Loss

Typical designs of electrodynamic actuators either have a winded conductor which in itself is self-supportive, or which is wound on a coil-carrier (fig. 9.4). The available space for the electrical coil within the homogenous magnetic flux is limited (A_{Coil}). The number of coil-turns $N_{\mathrm{Conductor}}$ is limited too within this area due to the cross-sectional surface a single turn needs $A_{Conductor}$. This cross-sectional surface is typically more than the actual cross-section of the conductor, as the winding will have gaps in between single turns (equ. 9.15). Additionally the actual conducting core with the cross sectional surface A_{Core} will be smaller than the cross-section of the conductor itself due to its isolation. Both parameters describing the geometrical losses in cross sections which are available within tables of technical handbooks [187] and are assumed as factors $k \geq 1$ according to equation 9.14. The length l of the conductor can be easily calculated by multiplication with the numbers of turns and the mean circumference *Circ* of the coil (equ. 9.16).

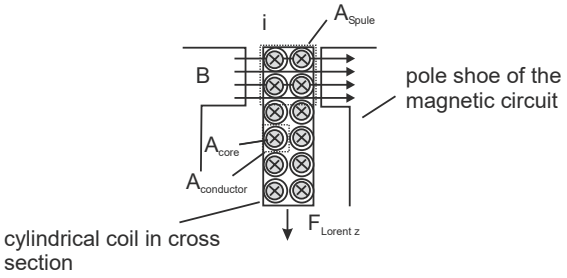


Fig. 9.4 Cross-section through a cylindric electrodynamic actuator according to the moving coil principle.

The choice of the conductor's diameter influences the resistance of the coil via the conducting area A_{Core} . The specific length-based resistance R_{spezf} of a conductor is given according to equation 9.17. Big conducting diameters with large cross-

sections A_{Core} allow coils able to conduct high currents at low voltages but - due to the limited volume available - few windings. Small diameters limit the necessary currents at high voltages and carry more windings. By a careful choice of wire-diameter the winding can be adjusted as a load to the corresponding source to drain the maximum available power.

The power loss P_{Loss} (equ. 9.18) acceptable within a given winding is limited. This limit is defined by the generated heat being able to dissipate. The technical solutions are dependent on the time of continuous operation, the thermal capacity resulting from the volume of the actuator and the materials it consists of, and a potential cooling system. A calculation of heat transmission is specific to the technical solution and can not be solved in general within this book. But the dependency of LORENTZ-force on power loss can be formulated:

$$A_{\mathrm{Conductor}} = k \cdot A_{\mathrm{Core}} \tag{9.14}$$

$$N_{\mathrm{Conductor}} = \frac{A_{\mathrm{Coil}}}{A_{\mathrm{Conductor}}} \tag{9.15}$$

$$l_{\mathrm{Conductor}} = N_{\mathrm{Conductor}} \cdot Circ \tag{9.16}$$

$$R_{\mathrm{spezf.}} = \frac{l_{\mathrm{Conductor}}\,\rho}{A_{\mathrm{Conductor}}} \tag{9.17}$$

$$P_{\mathrm{Loss}} = i^2 \cdot R_{\mathrm{Coil}} \tag{9.18}$$

From equation 9.18 follows

$$i = \sqrt{\frac{P_{\mathrm{Loss}}}{R_{\mathrm{Coil}}}} \tag{9.19}$$

With equation 9.17 there is

$$i = \sqrt{\frac{P_{\mathrm{Loss}}\,A_{\mathrm{Core}}}{\rho\,l_{\mathrm{Conductor}}}} \tag{9.20}$$

put into equation 9.9 (keeping the direction of current flow $\mathbf{e}_i$) there is

$$F_{\mathrm{Lorenz}} = \sqrt{\frac{P_{\mathrm{Loss}}\,A_{\mathrm{Core}}\,l_{\mathrm{Conductor}}}{\rho}}\,\mathbf{e}_i \times \mathbf{B} \tag{9.21}$$

by considering equation 9.15 to 9.16 the result is

$$F_{\mathrm{Lorenz}} = \sqrt{\frac{P_{\mathrm{Loss}}\,A_{\mathrm{Coil}}\,N\,\mathrm{Circ}}{\rho\,k}}\,\mathbf{e}_i \times \mathbf{B} \tag{9.22}$$

The equations 9.15 to 9.18 put into equation 9.9 gives a precise picture of the influence values on the LORENTZ-force (equ. 9.22). The level of Lorentz-force is given by the power loss P_{Loss} acceptable within the coil. If there is room for modifications to the geometrical design of the actuator, the cross-sectional area of the coil and the circumference of the winding should be maximized. Additional a choice of

alternative materials (e.g. alloy instead of copper) may minimize the electrical resistance. Furthermore the filling-factor k should be reduced. One approach could be the use of wires with a rectangular cross section to avoid empty spaces between the single turns.

The question for the maximum current itself is only relevant in combination with the voltage available and in the context of adjusting the electrical load to a specific electrical source. In this case for i_{Source} and u_{Source} the corresponding coil-resistance has to be chosen according to equation 9.24.

$$P_{\text{Source}} = u_{\text{Source}} \cdot i_{\text{Source}} = i_{\text{Source}}^2 \cdot R_{\text{Coil}} \tag{9.23}$$

$$R_{\text{Coil}} = \frac{P_{\text{Source}}}{i_{\text{Source}}^2} \tag{9.24}$$

Surprisingly from the perspective of a realistic design an increase in current is not necessarily the preferred option to increase the LORENTZ-force according to equation 9.22. The possibility to optimize P_{Loss} by adding cooling, or to analyze the temporal pattern of on- and off-times is much more relevant. Additionally the flux-density **B** has a - compared to all other influence factors - quadratically higher influence on the maximum force.

9.2.1.3 Maximization of the Magnetic Flux-Density B

For the optimization of electrodynamic actuators a maximization of the flux density **B** is necessary within the area where the conducting coils are located. This place is usually called air-gap and resembles an interruption of the otherwise closed ferromagnetic core conducting the magnetic flux. The heights of the magnetic flux density is influenced by

- the choice of ferromagnetic material for the magnetic core,
- the field winding/exciter winding of the static magnetic field, and
- the geometrical design of the magnetic core.

In the context of this book some basic design criteria for magnetic circuits are given. For an advanced discussion and optimization process source [125] is recommended.

Basics for the Calculation of Magnetic Circuits

Calculating magnetic circuits show several parallels to the calculation of electrical networks. As shown in table 9.2 several analogies between electrical and magnetic variables can be defined.

The direct analogy to the magnetic flux ϕ is the electrical charge Q. For the application of equations however it is helpful to regard electrical currents I as counterpart

Table 9.2 Analogies between electric and magnetic values.

Description	Electric	Magnetic
flux differential flux	charge Q [C=As] $I = \frac{dQ}{dt}$ [A]	magnetic flux ϕ [Vs]
flux-value	dielectric flux density D [C/m^2] $Q = \int_A \mathbf{D} d\mathbf{A}$ current density J [A/m^2] $I = \int_A \mathbf{J} d\mathbf{A}$	flux density B [T=Vs/m^2] $\phi = \int_A \mathbf{B} d\mathbf{A}$
El.-mag. coupling formerly:	voltage U [V] Electromotive force (e.m.f.)	flux/ampere turns Θ [A] Magnetomotive force (m.m.f.)
induction-laws	$U = -N\frac{d\phi}{dt}$	$\Theta = N\frac{dQ}{dt}$ $\Theta = NI$ (N= turns)
field-values	el. field strength E [V/m]	magn. field strength H [A/m]
differential-values	voltage U [V] $U = \int_a^b \mathbf{E} d\mathbf{s}$	magnetic voltage V [A] $V = \int_a^b \mathbf{H} d\mathbf{l}$
mesh-equations	$U_{\text{ges}} = \sum_i U_i$	$\Theta = \sum_i V_i$
resistances	el. resistance R [Ω] $R = \frac{U}{I}$	magn. resistance R_m [A/Vs] reluctance $R_m = \frac{V}{\phi}$
coupling factors	permittivity $\varepsilon = \varepsilon_0 \varepsilon_r$ ($\varepsilon_0 = 8,854 \cdot 10^{-12}$ C/Vm)	permeability $\mu = \mu_0 \mu_r$ ($\mu_0 = 1,256 \cdot 10^{-6}$ Vs/Am)
coupling between field- and flux-values	$\mathbf{D} = \varepsilon \mathbf{E}$	$\mathbf{B} = \mu \mathbf{H}$
power [W]	$P_{\text{el}} = U \cdot I$	
energie [J]	$W_{\text{el}} = W_{\text{el}} t$	$W_{\text{mag}} = \phi V$ $W_{\text{mag}} = \sum_n H_n l_n \cdot B_n A_n$

to the magnetic flux. Please note that this is an aid for thinking and not a mathematical reality, although it is very common. The actual direct analogy for the current I would be a time dependent magnetic flux $\frac{d\phi}{dt}$, which is usually not defined with an own variable name. The great exception with this model is the magnetomotive force Θ, which resembles the sum of all magnetic voltages V identical to a rotation within an electrical network. Nevertheless it is treated differently, as many applications require generating a magnetomotive force Θ to be generated by a certain number of winding-turns N and a current I, often referred to as ampere turns. The coupling between field- and flux-variables is given by the permittivity ε in case of electrical values and by the permeability μ in case of magnetic values. It is obvious

that the field-constants ε_0 differs from μ_0 by the factor 10^6. This is the main reason for the electromagnetic effect being the preferred physical realization of actuators in macroscopic systems[7].

However there is another speciality with the field-constants. The electrical permittivity can be regarded as quite constant (section 9.5) even for complex actuator designs, and can be approximated as linear around an operating point. The permeability μ_r of typical flux-conducting materials however shows a strong nonlinear relationship, the materials are reaching saturation. The level of magnetic flux has to be limited to prevent saturation-effects in the design of magnetic core.

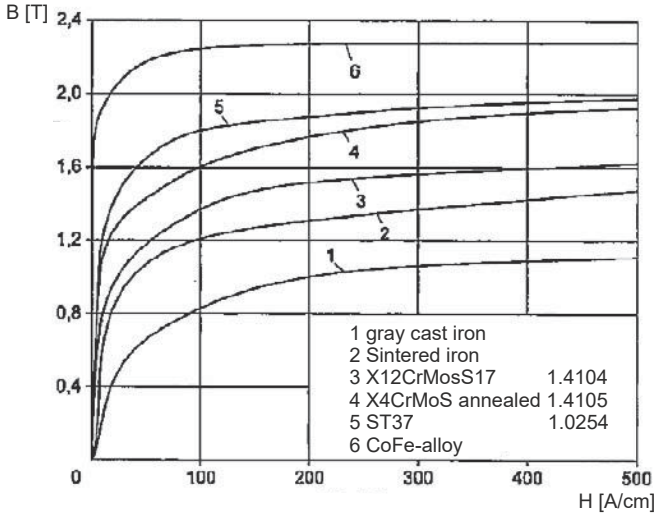


Fig. 9.5 Saturation curve of typical magnetic materials [125].

Magnetic Circuits

For the maximization of the magnetic flux density it is necessary to either analyze the magnetic circuit mathematical-analytically and/or do a numerical simulation of it. For the simulation of magnetic fields common CAD and FEM products are available[8]. For classification of the mathematical problem three solution levels exist:

[7] In micro-mechanical systems the energy-density relative to the volume becomes more important. The manufacture of miniaturized plates for capacitive actuators is much easier to realize with batch processes than the manufacture of miniaturized magnetic circuits.

[8] For the very beginning there are several free or open software-projects available for electrical and magnetic field simulation, e.g. for rotatory or planar systems a program from DAVID MEEKER named “FEMM”.

stationary, quasi-stationary, and dynamic magnetic fields. With stationary magnetic fields there is no time dependent change of the magnetic circuit. A steady state of flux density is assumed. With quasi-stationary field the induction is being considered resulting from changes within the current generating the magnetic field or a linearized change within the geometry of the magnetic circuit (e.g. a movement of an anchor). Dynamic magnetic fields consider additional effects covering the dynamic properties of moving mechanical components up to the change of the geometry of the magnetic circuit and the air gaps during operation. By dealing with electrodynamic actuators the analysis of static magnetic circuits is sufficient for a first dimensioning. The relevant dynamic drawbacks for electrodynamic actuators are presented in section 9.2.1.4.

There are two principle possibilities to generate the magnetic flux densities within the volume of the conducting coil:

1. Generation via winded conductors with another coil (exciter winding).
2. Generation via a permanent magnet.

Both approaches show specific pros and cons: With a winded conductor the flux density $B = \mu\,(N\,I - H_{Fe}\,l_{Fe})$ can be raised without any theoretical limit. In practical application the flux-conducting material will reach saturation (fig. 9.5) actually limiting the achievable maximum flux density. Additionally the ohmic resistance of the winding will generate electrical power losses, which will have to be dissipated in addition to the losses resulting from the electrodynamic principle itself (section 9.2.1.1). Abandoning any flux-conducting material and using exciter-windings with extremely low electrical resistance extraordinary high field-densities can be reached[9]. Till now such a technological effort for haptic devices is not made yet.

Building a magnetic circuit with a permanent magnet, the practical border for the flux density is given by the remanence flux density B_r of the magnetic material. Such a magnet can be compared to a source providing a certain magnetic power. The flux density - being the relevant quality for electrodynamic actuators - is not independent from the magnetic load attached to the permanent magnet. Additionally the relevant properties of the magnetic material are temperature-dependent, and wrong use of specific magnet-materials may harm its magnetic properties[10].

Nevertheless modern permanent-magnetic materials made of “rare earths” are the preferred source to generate static magnetic fields for electrodynamic actuators. The following section gives some basics on the calculation for simple magnetic circuits. In extension to what is shown here a more precise analytical calculation is possible [125]. However it is recommended to use simulation tools early within the design

[9] Magnet-resonance-tomographs for medical imaging generate field densities of 2 Tesla and more within air gaps of up to 1 m diameter by the use of supra-conducting coils and almost no magnetic circuit at all.

[10] e.g. when removing AlNiCo magnets out of their magnetic circuit after magnetization, they may drop below their coercive field strength actually losing performance.

process. Especially leakage fields are a great challenge for the design of magnetic circuits. And especially beginners should develop a feeling for the look of these fields with the aid of simulation tools.

Direct Current Magnetic Field

Figure 9.6a shows a magnetic circuit out of iron with a cross section A and an air-gap with the length ξ_G (G=Gap). The magnetic circuit has a winding with N turns conducting a current I. The medium length of the magnetic circuit is l_{Fe}. For calculation the circuit can be transformed into a magnetic equivalent network (fig. 9.6b). According to the analogies defined in table 9.2 the magnetic induction generates a magnetomotive force Θ as a differential value. In combination with two magnetic resistances of the iron circuit R_{mFe} and the air gap R_{mG} a magnetic flux ϕ can be identified.

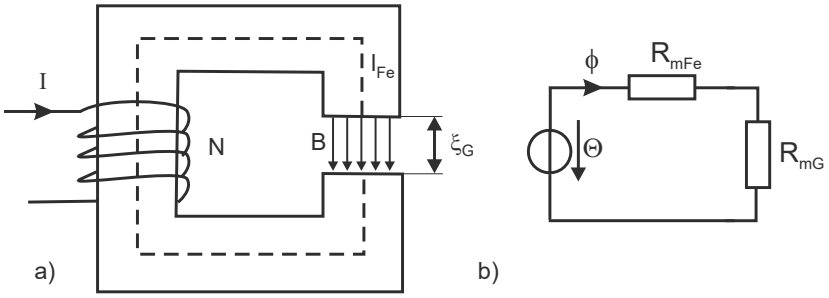


Fig. 9.6 Magnetic field generation B via a current-conducting coil with N turns (a), and derived equivalent circuit representation (b).

For the calculation of the flux density B in the air gap, it is assumed that this magnetic flux ϕ is identical to the flux within the iron part of the circuit. Leakage fields are disregarded in this example[11].

$$B = \frac{\phi}{A} \tag{9.25}$$

The magnetic resistance of materials and surfaces are dependent on the geometry and can be found in special tables [125]. For the magnetic resistance of a cylinder of the length l and the diameter d a resistance according to equation 9.26 is given.

[11] Considering leakage fields would be identical to a parallel connection of additional magnetic resistors to the resistance of the air gap.

$$R_m = \frac{4\,l}{\mu\,\pi\,d^2} \tag{9.26}$$

For the magnetic circuit the magnetic resistances $R_{m\mathrm{Fe}}$ and $R_{m\mathrm{G}}$ can be regarded as known or at least calculable. The magnetic flux is given by

$$\phi = \frac{\Theta}{R_{m\mathrm{Fe}} + R_{m\mathrm{G}}}, \tag{9.27}$$

and the flux density by

$$B = \frac{\Theta}{(R_{m\mathrm{Fe}} + R_{m\mathrm{G}})\,A}. \tag{9.28}$$

Using this procedure a clever approximation of the magnetic resistances of any complex network of magnetic circuits can be made. In this specific case of a simple horseshoe-formed magnet an alternative approach can be chosen. Assuming that the magnetic flux density in the air-gap is identical to the flux density in the iron (no leakage fields, see above) the flux-density B is given by:

$$B = \mu_0 \mu_r H \tag{9.29}$$

Assuming that μ_r is given either as a factor or with a characteristic curve (like in fig. 9.5) only the magnetomotive force Θ within the iron has to be calculated. With

$$\Theta = H_{\mathrm{Fe}}\,l_{\mathrm{Fe}} + H_{\mathrm{G}}\,\xi_G = \frac{B}{\mu_0\mu_r}\,l_{\mathrm{Fe}} + \frac{B}{\mu_0}\,\xi_G \tag{9.30}$$

the flux density

$$B = \Theta\,\frac{1}{\frac{l_{\mathrm{Fe}}}{\mu_0\mu_r} + \frac{\xi_G}{\mu_0}}, \tag{9.31}$$

results and can be written down immediately.

Permanent Magnets Generating the Magnetic Field

As stated earlier the typical approach to generate the magnetic field within an electrodynamic actuator is the choice of a permanent magnet. Permanent magnets can not simply be regarded as flux- or field-sources. Therefore some basic understanding of magnet technology will be necessary.

As a simple approach a magnet is a source of energy which is proportional to the volume of the magnet. Magnets are being made out of different magnetic materials (tab. 9.3) differing in the maximum achievable flux density (remanence flux density B_r), the maximum field-strengths (coercive field strength H_{cB} and H_{cJ})), and their energy density BH_{max}, such as the temperature coefficient. Additionally identical materials are differentiated according to being isotropic or anisotropic. With isotropic magnets its substance is made of homogeneous material which can be magnetized in one preferred direction. With anisotropic material a magnetic pow-

der was mixed with a binding material (e.g. epoxy) and formed via a casting or injection-molding process. Latter approach enables almost unlimited freedom for the magnet's geometry and a very large influence concerning the pole-distribution on the magnet. However anisotropic magnets are characterized by slightly worse characteristic values in energy density such as maximum field-strengths and flux densities.

Table 9.3 Magnetic properties of permanent-magnet materials [125]

Matrerial	B_r [T]	H_{cB} [kA/m]	$(BH)_{max}$ [kJ/m^3]
AlNiCo (isotrop)	0,5 ... 0.,9	10 ... 100	3 ... 20
AlNiCo (anisotrop)	0,8 ... 1,3	50 ... 150	30 ... 70
Hard ferrite (isotrop)	0,2 ... 0,25	120 ... 140	7 ... 9
Hard ferrite (anisotrop)	0,36 ... 0,41	170 ... 270	25 ... 32
SmCo (anisotrop)	0,8 ... 1,12	650 ... 820	160 ... 260
NdFeB (anisotrop)	1,0 ... 1,47	790 ... 1100	200 ... 415

Figure 9.7 shows the second quadrant of the *B*-*H*-characteristic curve (only this quadrant is relevant for an application of a magnet within an actuator) of different magnetic materials. The remanence flux density B_r equals the flux density with short-circuit pole shoes (a magnet being surrounded by ideal iron as magnetic circuit). When there is an air gap within the magnetic circuit (or even by the magnetic resistance of the real magnetic circuit material itself), a magnetic field strength H appears as a load. As a reaction an operation point is reached, which is shown here as an example on a curve of NdFeB for a flux-density of ≈200 kA/m. The actually available flux density at the poles is decreased accordingly. As electrodynamic actuators for haptic applications face high requirements according to their energy density, there are almost no alternatives to the usage of magnet materials based on rare earths (NdFeB, SmCo). This is very accommodating for the design of the magnetic circuit, as nonlinear effects near the coercive field strength such as with AlNiCo or Barium-ferrite are of no relevance[12]. Rare earth magnets allow an approximation of their B/H-curve with a linear equation, providing a very nice relationship for their magnetic resistance (fig. 9.8c):

$$R_{\mathrm{Mag}} = \frac{V}{\phi} = \frac{H_c\, l_{\mathrm{Mag}}}{B_r\, A} \tag{9.32}$$

[12] The small coercive field strength of these materials e.g. result in the effect, that a magnet magnetized within a magnetic circuit does not reach its flux density anymore once removed and even after re-assembly into the circuit again. This happens due to the temporary increase of the air gap, which is identical to an increase of the magnetic load to the magnet beyond the coercive field strength. Additionally the temperature-dependency of the coercive field strength and of the remanence flux density is critical. Temperatures just below the freezing point may result in a demagnetization of the magnet.

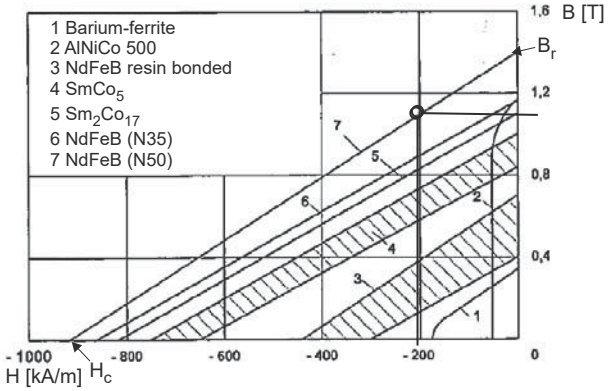


Fig. 9.7 Demagnetization curves of different permanent-magnet materials [125].

With this knowledge the magnetic circuit of figure 9.8a and the corresponding equivalent circuit (fig. 9.8b) can be calculated identical to an electrically excited magnetic circuit.

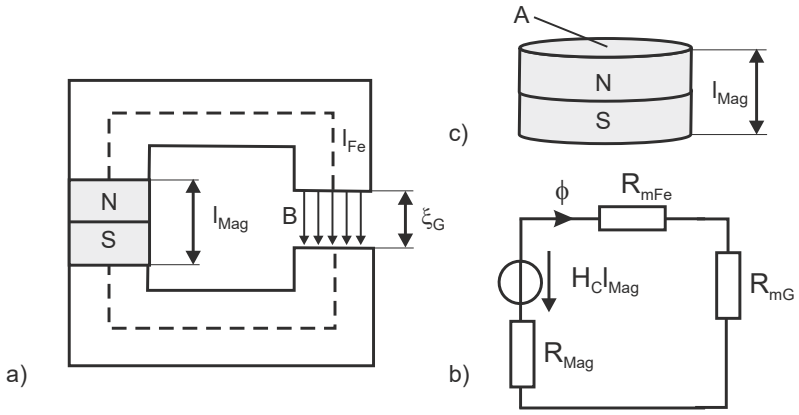


Fig. 9.8 Magnetic field generation *B* via permanent magnets (a), derived equivalent circuit (b), and dimensions of the magnet (c).

The flux density within the iron is once again given by

$$B = \frac{\phi}{A} \tag{9.33}$$

For the given magnetic circuit the resistances $R_{m\mathrm{Fe}}$ and $R_{m\mathrm{G}}$ are assumed as known or calculable. From equation 9.32 the magnetic resistance of the permanent

magnet is known. The source within the equivalent circuit is defined by the coercive field strength and the length of the magnets $H_c\, l_{\mathrm{Mag}}$. These considerations result in

$$\phi = \frac{H_c\, l_{\mathrm{Mag}}}{R_{m\mathrm{Fe}} + R_{m\mathrm{G}} + R_{\mathrm{Mag}}}, \tag{9.34}$$

and the flux density

$$B = \frac{H_c\, l_{\mathrm{Mag}}}{(R_{m\mathrm{Fe}} + R_{m\mathrm{G}} + R_{\mathrm{Mag}})\, A}. \tag{9.35}$$

Slightly rearranged and R_{Mag} included gives

$$B = \frac{B_r H_c \frac{l_{\mathrm{Mag}}}{A}}{(R_{m\mathrm{Fe}} + R_{m\mathrm{G}})\, B_R + H_c \frac{l_{\mathrm{Mag}}}{A}}. \tag{9.36}$$

Equation 9.36 states by the factor $B_r H_c \frac{l_{\mathrm{Mag}}}{A}$ that it is frequently very helpful for achieving a maximum flux density B in the air gap to increase the length of a magnet with at the same time minimized cross-sectional area of the magnetic circuit - which is of course limited by the working distance within the air gap and the saturation field strengths of the magnetic circuit.

9.2.1.4 Additional Effects in Electrodynamic Actuator

To do a complete characterization of an electrodynamic actuator there are at least three more effects, whose influences will be sketched within the following paragraphs.

Induction

For a complete description of an electrodynamic actuator beside the geometrical design of its magnetic circuit and the mechanical design of its winding and the considerations concerning the electrical power losses, its other dynamic electrical properties have to be considered. For this analysis the electrodynamic actuator is regarded as a bipolar transformator (fig. 9.9).

A current $\underline{i}_0$ generates via the proportional constant $B\,l$ a force $\underline{F}_0$, which moves the mechanical loads attached to the actuator. The movement itself results in a velocity $\underline{v}_0$ which is transformed via the induction law and the proportional constant to an induced voltage $\underline{u}_1$. By measurement of $\underline{u}_1$ and a current source the rotational velocity or the movement velocity v can be measured, with a voltage source the measurement of $\underline{i}_0$ provides a force- or torque-proportional signal. This is the approach taken by the variant of admittance controlled devices as a control value (see section 5.1.4).

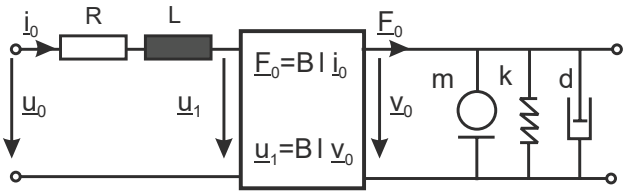


Fig. 9.9 Electrical and mechanical equivalent circuit of an electrodynamic actuator as being a transformer.

The induction itself is a measurable effect, but should not be overestimated. Typically electrodynamic actuators are used within haptic systems as direct drives at small rotational or translational velocities. Typical coupling factors with rotatory drives are - depending on the size of the actuators - in an area between 100 to 10 $\frac{revolutions}{sV}$. At a rotational speed which is already fast for direct drives of 10 Hz, induced voltage amplitude $|\underline{u}_1|$ of 0.1 to 1 V can be achieved. This is around 1% to 5% of the control voltage's amplitude.

Electrical Time Constant

Another aspect resulting from the model according to figure 9.9 is the electrical transfer characteristics. Typical inductances L of electrodynamic actuators lie in the area of 0.1 mH to 2 mH. The ohmic resistance of the windings is largely depending on the actual design, but as a rule of thumb values between $10\,\Omega$ to $100\,\Omega$ can be assumed. The step-response of the electrical transfer system $\frac{\underline{i}_0}{\underline{u}_0}$ shows a time-constant $\tau = \frac{L}{R} = 10\,\mu s$ to $30\,\mu s$ and lies within a frequency range »10 kHz, which is clearly above the relevant dynamic area of haptics.

Field Response

A factor which can not so easily be neglected when using electrodynamic actuators for high forces is the feedback of the magnetic field generated by the electromagnetic winding on the static magnetic field. Taking the actuator from the example at the beginning (fig. 9.3 on page 200) positive currents generate a field of opposite direction to the field generated by the magnet. This influence can be considered by substitution of both field sources. Depending on the direction of current this field either enforces or weakens the static field. With awkward dimensioning this can result in a directional variance of the actuator properties. The problem is not the potential damage to the magnet, modern magnetic materials are sufficiently stable, but a variation of the magnetic flux density available within the air gap. An intended application of this effect within an actuator can be found in an example according

to figure 9.29 on page 235.

A deeper discussion about electrodynamic actuators based on concentrated elements can be found in [158].

9.2.2 Actual Actuator Design

As stated earlier electrodynamic actuators are composed of three basic components: coil/winding, magnetic circuit, and magnetic exciter. The following section describes a procedure for the design of electrodynamic actuators based on these basic components. As the common principle for excitation a permanent magnet is assumed.

9.2.2.1 Actuator Topology

The most fundamental question for the design of an electrodynamic actuator is its topology. Usually it is known whether the system shall perform rotary or translatory movements. Afterward the components magnetic circuit, the location of magnets, pole-shoes and the coil itself can be varied systematically. A few quite common structures are shown in figure 9.11 for translational actuators, and in figure 9.10 for rotatory actuators. For the design of electrodynamic actuators in any case the question should be asked, whether the coil or the magnetic circuit shall move. By this variation apparently complex geometrical arrangements can be simplified drastically. Anyway it has to be considered that a moving magnet has more mass and can typically be moved less dynamically than a coil. On the other hand there is no contact- or commutating problem to be solved with non-moving windings.

Moving Coils

Electrodynamic actuators according to the principle of moving coils with a fixed magnetic circuit are named “moving coil” in the case of a linear movement and “iron-less rotor” in the case of a rotatory actuator. They always combine few moving masses and as a result high dynamics. The translatory version shows displacements of a few millimeters, and is used especially within audio applications as loudspeaker. Actuators according to the principle of “moving coils” have two disadvantages:

- As the coil is moving, the electrical contact is subject to mechanical stresses. Especially with high displacements the contact has to be mechanically robust.
- If there is the idea to design moving coils as pure force sources with large displacements, always only a small area of the conducting coil is within the air-gap and therefore contributes to the force generation. With large displacements

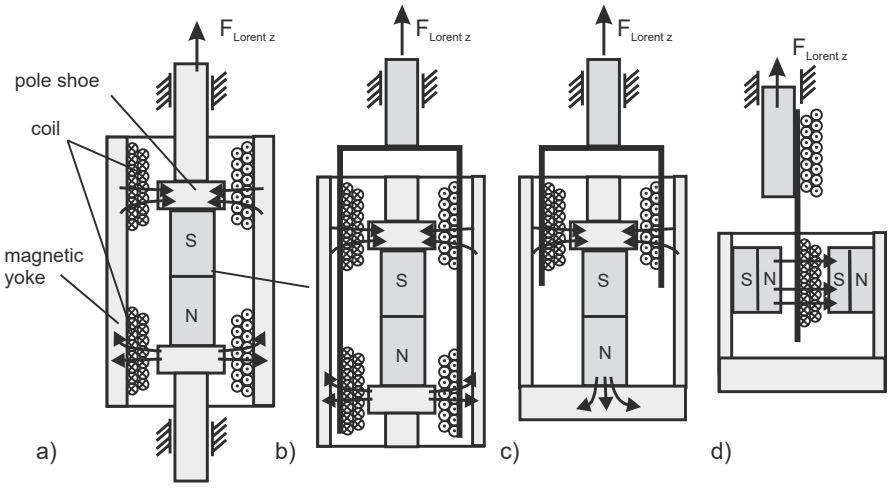


Fig. 9.10 Variants of electrodynamic actuators for translational movement with moving magnets (a), moving coils (b), as plunger-type (c), and as flat-coil (d).

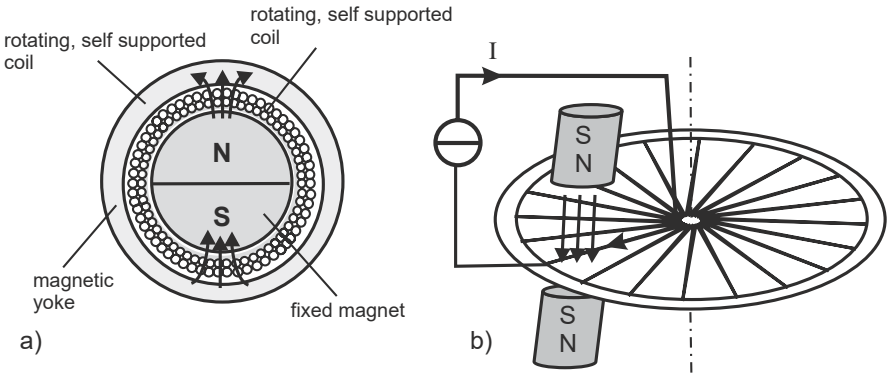


Fig. 9.11 Variants of electrodynamic actuators for rotatory movements with self-supportive winding (a), and with disc-winding.

moving coils show an even lower efficiency factor. This can be compensated by switching the active coil areas, which again results in the necessity to have more contacts.

A similar situation happens with rotatory systems. Based on the electrodynamic principle there are two types of windings applicable to rotatory servo-systems: the FAULHABER and the MAXON-winding of the manufacturers with identical names. These actuators are also known as "iron-less" motors. Both winding principles allow the manufacture of self-supportive coils. A diagonal placement of conductors and a baking process after winding generates a structure sufficiently stable for the

centrifugal forces during operation. The baked coils are connected with the rotating axis via a disk. The complete rotor (fig. 9.12) is build of these three components. By the very small inertia of the rotor such actuators show impressive dynamic properties. The geometrical design allows placing the tubular winding around a fixed diametral-magnetized magnet. This enables another volume reduction compared to conventional actuators as its housing has to close the magnetic circuit only instead of providing additional space for magnets.

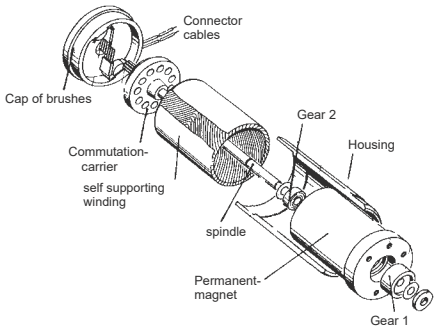


Fig. 9.12 Design of an electrodynamic actuator with self-supportive winding according to the FAULHABER-principle [97].

Within the self-supportive winding there are areas of parallel lying conductors combined to poles[13]. With moving coils there is always the need for a specialized contactor, either via contact rings, or electronic commutation or via mechanical switching. Depending on the number of poles all coils are contacted at several points. In case of mechanical switching these contacts are placed on the axis of the rotor and connected via brushes with the fixed part of the actuator named “stator”. This design enables a continuous movement of the rotor, whereas a change of the current flow is made purely mechanically by the sliding of the brushes on the contact areas of the poles on the axis. This mechanical commutation is a switching procedure with an inductance placed in parallel.
As such an actuator can be connected directly to a direct current source, they are known as “DC-drives”. As stated within section 9.1 the term “DC-drive” is not only limited to actuators according to the electrodynamic principle but is also frequently applied to actuators following the electromagnetic principle (section 9.3).

[13] The FAULHABER and the MAXON excel by a very clever winding technique. On a rotating cylinder respectively a flatly pressed rectangular winding poles can be combined by contacting closely located areas of an otherwise continuous wire.

Moving Magnet

In case of translatory (fig. 9.10a) systems actuators according to the principle of a moving magnet are designed to provide large displacements with compact windings. The moving part of the actuator is composed almost completely of magnetic material. The polarity direction of this material may vary in its exact orientation. Actuators according to this principle are able to provide large power, but are expensive due to the quantity of magnet material necessary. Additionally the moving magnet is heavy; the dynamics of the actuator is therefore smaller than in the case of a moving coil.
In case of a rotatory system a design with moving magnet is comparable to a design with moving coil. Figure 9.13 shows such a drive. The windings fixed to the stator are placed around a diametral magnetized magnet. It rotates on an axis, which frequently additionally moves the magnetic circuit too. Providing the right current feed to the coil the orientation of the rotor has to be measured. For this purpose sensors based on the Hall-effect or optical code wheels are used.

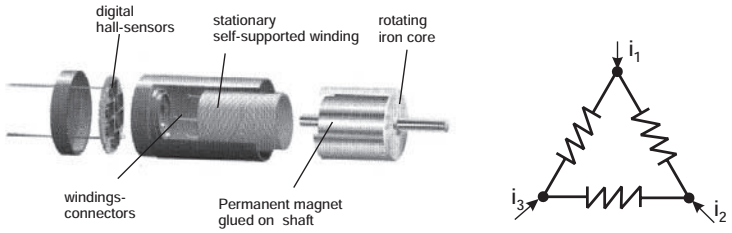


Fig. 9.13 Components of a EC-drive and equivalent circuit.

Electrodynamic actuators with moving magnet are known as EC-drives (electronic-commutated). This term is not exclusive to electrodynamic actuators, as there are electronic-commutated electromagnetic drives too. EC-drives - whether they are electrodynamic or electromagnetic - combined with the corresponding driver electronics are frequently known as servo-drives. Typically a servo-drive is an actuator able to follow a predefined movement path. Servo-drives are rarely used for haptic devices. However the usage of EC-drives for haptic application is very frequent, but then they are equipped with specialized driver electronics.

9.2.2.2 Commutation in the Context of Haptic Systems

The necessary commutation of the current for rotating actuators has a big influence upon the quality of force- respective torque-output.

Mechanically Commutating Actuators

With mechanically commutating actuators the current flow is interrupted suddenly. Two effects of switching contacts appear: The voltage at the contact point increases, sparks may become visible - an effect which is called electrical brush sparking. Additionally the remaining current flow induces a current within the switched-off part of the winding which itself results in a measurable torque. Figure 9.14b shows a measurement of a force acting upon a rod driven by a mechanically commutated actuator according to figure 9.14a. The force ripples excite the base force of 1.3 N by$\approx$ 100 mN, which is around 8%. A transformation of the measurement into the perceptional space $\tilde{K}$ according to the concept of quantification haptic perception (section 4.3) amplifies the difference even further. The perception-peaks resulting from commutation exceed the continuous perception by 25%. This amplification is a result from the lower frequencies being damped whereas higher frequency components are enhanced. The high frequencies of the commutating flanks dominate disproportionately high haptic perception.

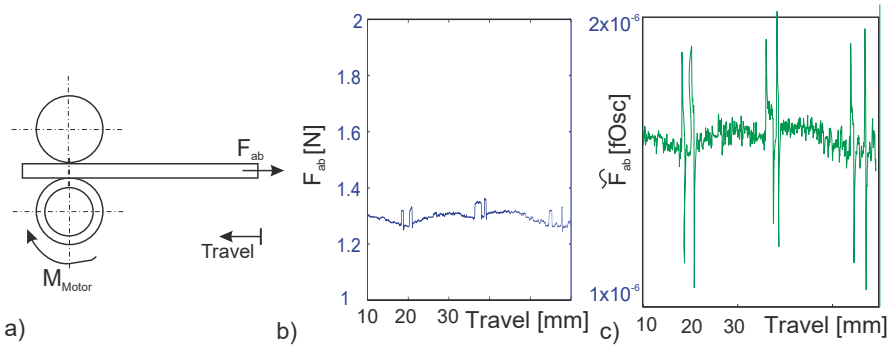


Fig. 9.14 Measured and perceived force of a mechanically commutated actuator.
a) Measurement setup, b) measured force plot, c) perceived force according to the *FIP*-method.

The current- and torque changes can be reduced by the inclusion of resistors and capacitors into the coil. However this results into high masses of the rotor and worse dynamic properties. Beside that a full compensation is impossible. Nevertheless mechanically commutating actuators are in use for inexpensive haptic systems. The Phantom Omni from SensAble and the Falcon from Novint use such actuators.

Electronic Commutated Electrodynamic Actuators

Electronic commutated electrodynamic actuators differ from mechanically commutated actuators by the measurement technology used as a basis for switching currents. There are four typical designs for this technology:

- In sensor-less designs (fig. 9.15a) an induced voltage is measured within a coil. At zero-crossing point one pole is excited with a voltage after an interpolated 30° phase delay dependent on the actual revolution speed of the rotor. In combination of measurement of the inductance followed by a switched voltage, a continuous rotation with batch-wise excitation is realized. This procedure can not be applied to low rotation speeds, as the induced voltage becomes too low and accordingly the switching point can hardly be interpolated. Additionally the concept of using one to two coils for torque generation results in a high torque variations at the output of up to 20%, making this approach not useful for haptic systems.
- Block-commutating procedures (fig. 9.15b) are based on the usage of simple hall-switches or field-plates for position detection of the rotor. Three sensors located at 120° angular phase shift allow the detection of six different rotor positions. Reducing positioning information to six orientations per revolution makes this approach equally inappropriate for haptic applications, as the torque varies in a range of >15% for one revolution.
- Sinus-commutating procedures with analogue hall-sensors are based on the measurement of the rotor position by at least two sensors. They are placed with an angle of either 120° or 90° at the front of the rotor. They provide voltages in an angular phase shift according to their geometrical position. By analyzing the polarity and the absolute height of the voltages absolute positioning information can be obtained and used for commutating the windings. If the phase lag between both sensor signals is identical to the phase lag between the poles of the winding a direct control of current-drivers can be performed without the need for a digitalisation or a specific calculation step.
- Sinus-commutating with digital code-wheels (fig. 9.15c) are based on the measurement of rotor position by the use of - usually optical - code discs. By reflective or transmittive measurement the rotor position is sampled with high resolution (e.g. 128 positions / turn). This relative positioning information can be used for position measurement after an initial calibration. Depending on the code-wheels resolution a very smooth sinoid commutation can be achieved with this method.

The sinus-commutating methods are the preferred solutions used for haptic applications due to the little torque variations and their applicability for slow revolution speeds typical to direct drives.

9.2.3 Actuator Electronics

Electrodynamic actuators require some specific electrical circuits. In the following section the general requirements on these electronics are formulated.

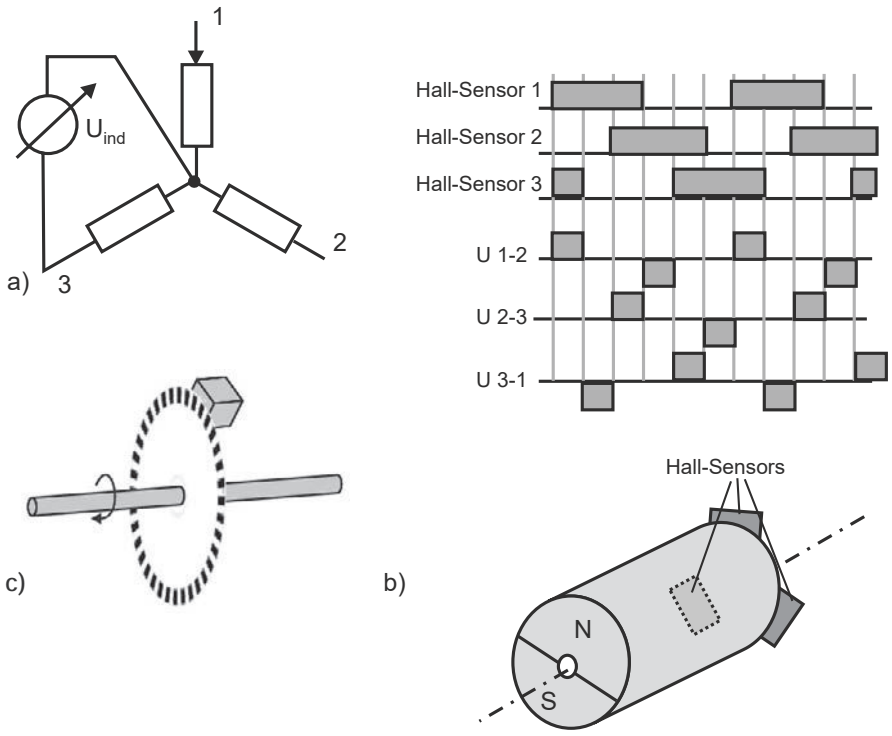


Fig. 9.15 Technologies for different commutation methods: sensor-less (a), block-commutation (b) and optical code-wheel (c)

9.2.3.1 Driver Electronics

Driver electronics are electrical circuits transforming a signal of low power (several volts, some milli-ampere) into a voltage- or current level appropriate to drive an actuator. For electrodynamic actuators in haptic applications driver electronics have to provide a current in a dynamic range from static up to several kilohertz. This paragraph describes general concepts and approaches for such circuits.

Topology of Electric Sources

Driver electronics for actuators - independently from the actuation principle they are used for - are classified according to the flow of electrical energy (fig. 9.16). There are four classes of driver electronics:

- 1-quadrant controllers are capable of generating positive output currents and voltages. An actuator driven by them is able to move in one direction. These controller use only the first quadrant according to figure 9.16a.
- Switched 1-quadrant controllers are capable of a direction change by the input of a logical signal. They are working within the 1st and 3rd quadrant according to figure 9.16a. The switching point is a nonlinear step in their characteristic curve.
- Real 2-quadrant controllers are capable of providing a characteristic curve which is steady around the zero point. They function in the 1st and 3rd quadrant according to figure 9.16a, but are not capable to conduct currents and voltages with opposite directions.
- 4-quadrant controllers function within all four quadrants of figure 9.16a. They are able to control currents and voltages in any combination of directions. Four-quadrant controllers allow energy recovery by induced currents to an energy storage, which is especially relevant for mobile applications.

For haptic application the switched 1-quadrant controller is frequently met, as many haptic systems do not have the necessity to control the device near the voltage- or current-zero point. However for systems with high dynamics and low impedance the 2-quadrant and the 4-quadrant controller are relevant, as the unsteadiness near the zero-point is perceivable with high quality applications.

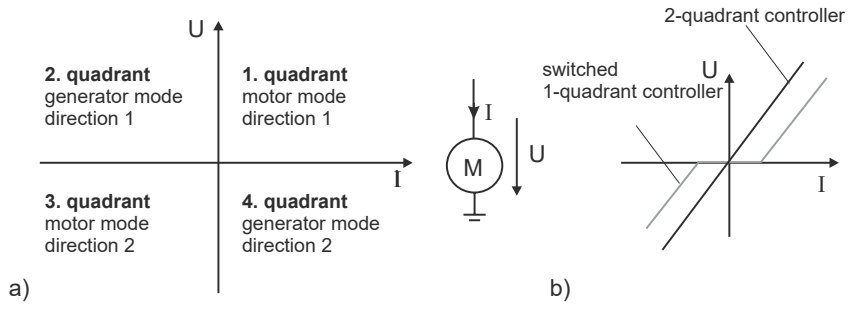


Fig. 9.16 Visualization of the four quadrants of an electric driver, formed by the directions of current and voltage

Pulse-Width-Modulation and H-Bridges

With the exception of some telemanipulators, the sources controlling the actuators are always digital processors. As actuators need an analogue voltage or current to generate forces and torques some transformer between digital signals and analogue control value is necessary. There are two typical realizations of these transformers:

1. Usage of a digital-analog converter (D/A-converter)
2. Usage of a pulse-width-modulation (PWM)

The use of D/A-converters as external components or integrated within a microcontroller is not covered further in this book, as it is, if necessary to use, extremely simple. It just requires some additional efforts in circuit layout. Latter results in it being not used much for the control of actuators.

With electrodynamic actuators the method of choice are driver electronics based on PWM (fig. 9.17a). With the PWM a digital output of a controller is switched with a high frequency (>10 kHz[14]). The period of the PWM is given by the frequency. The program controls the duty cycle between on- and off-times. Typically one byte is available to provide a resolution of 256 steps within this period. After filtering the PWM, either via an electrical low-pass or via the mechanical transfer-characteristics of an actuator, a smoothed output signal becomes available.

Pulse-width-modulation is frequently used in the combination with H-bridges (fig. 9.17b). The term H-bridge results from the H-like shape of the motor surrounded by four switches. The H-bridge provides two operation modes for two directions of movement and two operation modes for braking. If according to figure 9.17b the two switches S2 and S5 are on, the current I will flow through the motor in positive direction. If instead switches S3 and S4 are switched on, the current I will flow through the motor in negative direction. One additional digital signal acting upon the H-bridge will change the direction of movement of the motor. This is the typical procedure with switched 1-quadrant controllers. Additional switching-states are given by switching the groups S2 and S3 respectively S4 and S5. Both states results in short-circuit of the actuator and stops its movement. Other states like simultaneously switching S2 and S4 respectively S3 and S5 results in short-circuit of the supply voltage, typically destroying the integrated circuit of the driver. To combine the H-bridge with a PWM either switch-groups S2 and S5 can be switched according to the timing of the PWM, or additional switches S1 (fig. 9.17c) can be placed in series to the H-bridge modulating the supply voltage U. In a practical realization latter is the preferred design, as the timing of the switches S2 to S5 is very critical to prevent likely short circuits of the supply voltage. The effort to perform this timing between the switching is usually higher than the costs of another switch in series. The practical realization of H-bridges is done via field-effect transistors. The discrete design of H-bridges is possible, but not easy. Especially the timing between switching events, the prevention of short-circuits, and the protection of the electronics against induced currents is not trivial. There are numerous

[14] Typical frequencies lie in between 20 kHz to 50 kHz. However especially within automotive technology for driving LEDs, PWMs for current drivers with frequencies below 1 kHz are in application. Frequencies within this range are not applicable to haptic devices, as the switching in the control value may be transmitted by the actuator and will therefore be perceivable especially in static conditions. Typical device designs show mechanical low-pass characteristics even at frequencies in the area of 200 Hz already. However due to the sensitivity of tactile perception in an area of 100 to 200 Hz, increased attention has to be paid on any switched signal within the transmission chain.

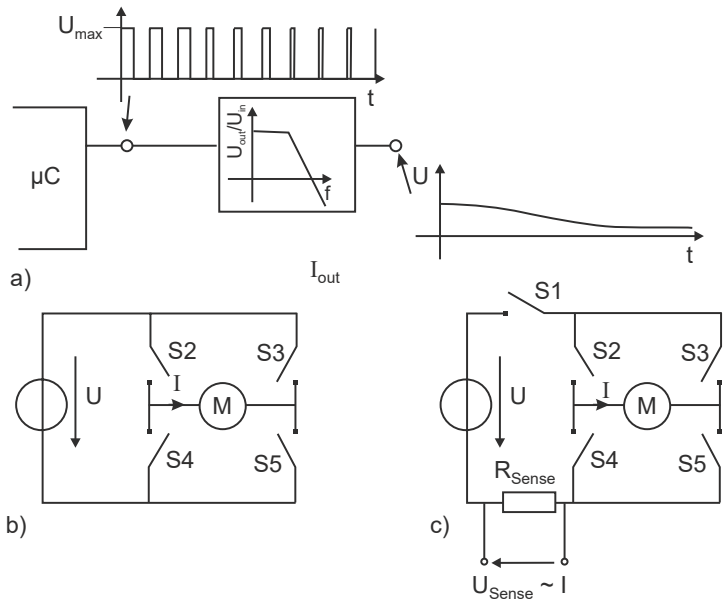


Fig. 9.17 Principle of puls-width-modulation (PWM) at a digital μC-output (a), h-bride circuit principle (b), and extended h-bridge with PWM (S1) and current measurement at (R_{Sense})(c)

integrated circuits available at the market which already include appropriate protective circuitry and provide only a minimum of necessary control lines. The ICs L6205 (2A), L293 (2.8A) and VNH 35P30 (30A) are some examples common with test-bed developments. For EC drives there are specific ICs performing the timing for the field-effect transistors and reducing the number of necessary PWMs from the microcontroller. The IR213xx series switches three channels with one external half-bridge per channel built up from N-MOS transistors with a secure timing for the switching events.

The PWM described above with an H-bridge equals a controlled voltage source. For electrodynamic systems such a control is frequently sufficient to generate an acceptable haptic perception. Nevertheless for highly dynamic haptic systems a counter induction (section 9.2.1.4) due to movement has to be expected, resulting in a variation of the current within the coils generating an uncontrolled change of the LORENTZ-force. Additionally the power-loss within the coils (paragraph 9.2.1.1) may increase the actuator's internal temperature resulting in a change of conductivity of the conductor's material. The increasing resistance with increasing temperatures of the conductor results in less current flow at a constant voltage source. An electrodynamic actuator made of copper as conductive material would generate a reduced force when operated. With higher requirements on the quality of haptic output a controlled current should be considered. In case of a PWM a resistor with low re-

sistance (R_{Sense} in fig. 9.17c) has to be integrated, generating a current-proportional voltage U_{Sense}, which itself can be measured with an A/D input of the controller. The control circuit is closed within the microcontroller. However the A/D transformation and the closing of the control circuit can be challenging for state of the art electronics with highly-dynamic systems with border frequencies of some kilohertz. Therefore analog circuits should be considered for closed-loop current controls too.

Analogue Current Sources

Analogue current sources are - to make it simple - controlled resistors within the current path of the actuator. Their resistance is dynamically adjusted to provide the wished current flow. Identical to classical resistors analogue current sources transform the energy which is not used within the actuator into heat. Consequently in comparison to the switched H-bridges they are generating a lot of power loss. By the use of a discrete current control (fig. 9.18a) analogue current sources for almost any output currents can be built by the choice of one to two field-effect-transistor (FET). For heat dissipation they are required to be attached to adequate cooling elements. There are only little requirements on the operational amplifiers themself. They control the FET within its linear range proportional to the current-proportional-voltage generated at R_{Sense}. Depending on the quadrant used within operational mode (1 or 3) either the N-MOS transistors or the P-MOS transistor is conductive. An alternative to such discrete designs is the use of power-amplifiers (e.g. LM675, fig. 9.18b). It contains fewer components and is therefore less dangerous to make errors. Realized as non-inverting or inverting operational amplifier with a resistor for measurement R_{Sense}, they can be regarded as a voltage controlled current source.

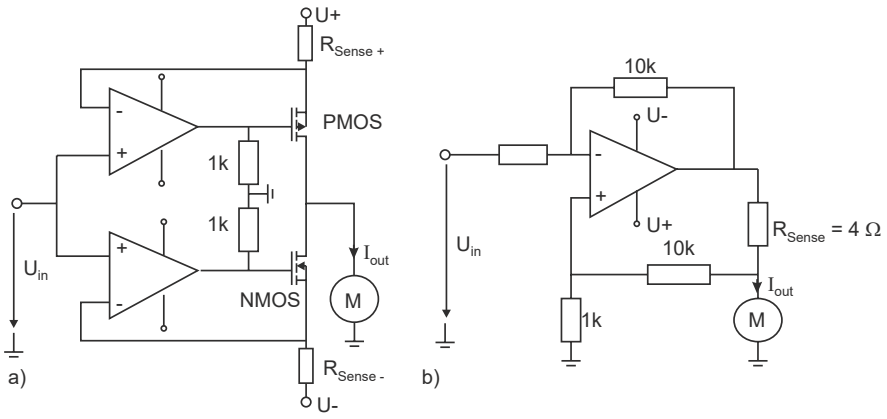


Fig. 9.18 Discrete closed-loop current control [255] (a), and closed-loop current control with a power-operational-amplifier (b).

9.2.3.2 Monitoring Temperature

Resulting from the low efficiency factor and the high dissipative energy from electrodynamic actuators it is useful to monitor the temperature nearby the coils. Instead of including a measuring resistor PT100 nearby the coil, another approach monitors the electrical resistance of the windings themselves. Depending on the material of the windings (e.g. cooper, Cu) the conductivity changes proportional to the coil's temperature. With copper this factor is 0.39% per Kelvin temperature change. As any driver electronics either works with a known and controlled voltage or current, measurement of the other component immediately provides all information to calculate resistance and consequently the actual coil temperature.

9.2.4 Examples for Electrodynamic Actuators in Haptic Devices

Electrodynamic actuators are most frequently used as force and torque sources within kinaesthetic systems. Especially EC-drives can be found in the products of Quanser, ForceDimensions, Immersion, and SensAble. Mechanically commutated electrodynamic actuators are used within less expensive devices, like the Phantom Omni or the Novint Falcon. For tactile applications electrodynamic actuators appear only with linear moving coils or magnets as oscillation sources. The possibility to control frequency and amplitude independently from each other makes them interesting for tactors. Tactors (fig. 9.19) are small, disk-like actuators which can be integrated e.g. in clothes or mobile devices to transmit information via tactile stimulation in small areas of the skin.



Fig. 9.19 Electrodynamic tactor by Audiological Engineering Inc., with a frequency range from 100 Hz to 800 Hz.

9.2.4.1 Cross-Coil System

Beside self supportive coils electrodynamic actuators according the design of cross coils are one possibility to generate defined torques. Continental VDO developed a haptic rotary actuator device being a central control element for automotive applications (fig. 9.20). It contains a diametral magnetized NdFeB-magnet. The magnet is surrounded by a magnetic circuit. The field-lines reach from the magnet to the magnetic circuit. The coils surround the magnet in an angular phase of 90°, and the electrodynamic active winding section lies in the air-gap between magnetic circuit and magnet. The rotary position control is made via two hall-sensors placed in a 90° position. The actuator is able to generate a ripple-free torque of ≈25 mNm at a geometrical diameter of 50 mm, which is additionally increased by an attached gear to ≈100 mNm torque output.

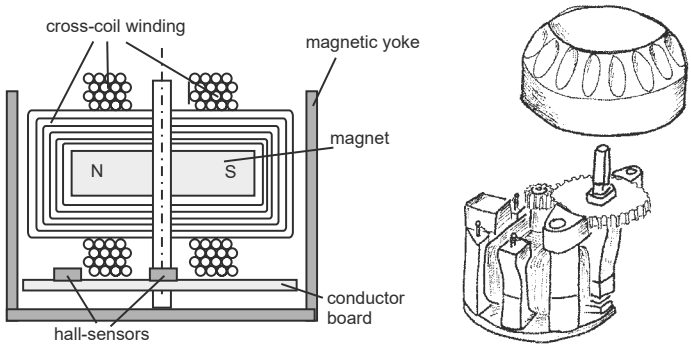


Fig. 9.20 Electrodynamic cross-coil system with moving magnet as haptic rotary actuator.

9.2.4.2 HANNAFORD- Flat-Coil System

The flat coil system (fig. 9.21) designed from the workgroup of HANNAFORD is part of the fingertip-haptic-display. The alloy wire (reduced specific resistance compared to copper) is wound in a coil in one plane and fixated by a backing process. For better conductance of the thermal loss it is integrated in a carrier made of alloy, being pivot in one axis. The magnetic circuit is closed above two blades of the coil with a diametral magnetized thin magnet. The system is able to generate torques of 160 mNm and more. The power management to dissipate the thermal loss and the analysis for convection is detailed in [160].

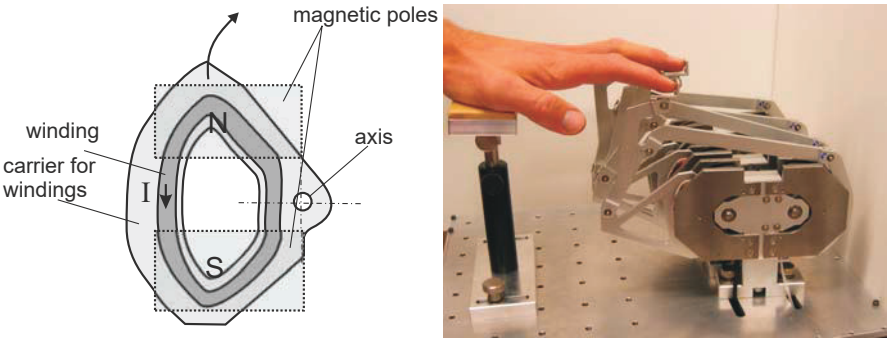


Fig. 9.21 Electrodynamic flat-coil system for the generation of forces a the finger-tips[160].

9.2.4.3 KUNSTMANN-Telemanipulator

The electrodynamic haptic device of the KUNSTMANN telemanipulator is part of a micro-assembly workplace. It aims at generating forces to two finger tips, and provides one additional force for gravity generation. A solution (fig. 9.22) of two rectangular wounded flat coils moving in an iron brace was chosen. Via diametral magnetized thin magnets a field normal to the brace is generated. Therefore the long sides of the coils are active for force generation, as they lie within the air gap. Forces up to 3 N can be generated by this device in continuous operation.

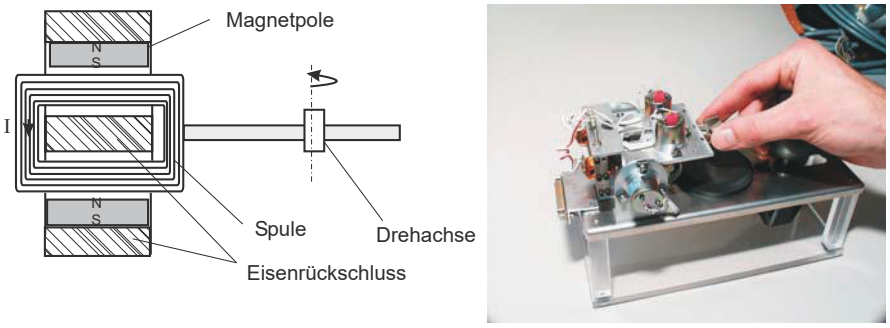


Fig. 9.22 Electrodynamic rectangular flat-coil system similar to a moving coil [147].

9.2.4.4 DOERRER-HapKeys

The HapKeys from DOERRER have been mentioned earlier in section 2.2.5.2 . The electrodynamic linear actuators building the basis of this device are equipped with

friction type bearings, and moving magnets with pole-shoes within cylindrically wound fixed coils. The coils have an inner diameter of 5.5 mm and an outer diameter of 8 mm. The magnetic circuit is decoupled from other nearby elements within the actuator-array. It is made of a tube with a wall thickness of 0.7 mm of a cobalt-iron alloy with very high saturation flux density. Each actuator is able to generate 1 N in continuous operation mode.

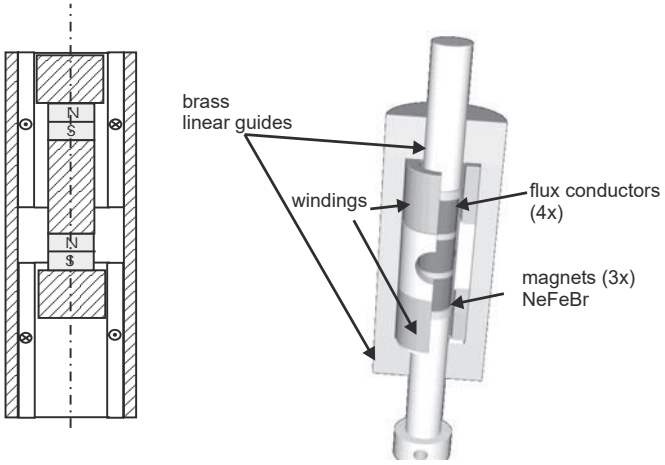


Fig. 9.23 Electrodynamic linear actuator with moving magnet [46].

9.2.5 Conclusion about the Design of Electrodynamic Actuators

Electrodynamic actuators are the preferred actuators used for kinaesthetic impedance-controlled haptic devices due to their proportional correlation between the control value “current” and the output-values “force” or “torque”. The market of DC -and EC-drives offers a wide variety of solutions, making it possible to find a good compromise between haptic quality and price for many applications. If there are special requirements to be fulfilled, the design, development, and start of operation of special electrodynamic actuator variants are quite easy. The challenges by thermal and magnetic design are manageable, as long as some basic considerations are not forgotten. The examples of special haptic systems seen in the preceding section prove this impressively. Just driver electronics applicable to haptic systems and its requirements are still an exceptional component within the catalogs of manufactures from automation-technology. They must either be paid expensively or be built by oneself. Therefore commercial manufacturers of haptic devices, e.g. Quanser, offer

their haptic-applicable driver electronics independent from the own systems for sale.

For the design of low-impedance haptic systems currently no alternative to electrodynamic systems is at hand. Other actuation principles which are discussed within this book need a closed-loop control to overcome their inner friction and nonlinear force/torque-transmission. This always requires some kind of measurement technology such as additional sensors or the measurement of inner actuator states. The efforts connected with this are still a big advantage for electrodynamic actuators, which is gained by a low efficiency factor and as a consequence the relatively low energy density per actuator-volume.

9.3 Electromagnetic Actuators

THORSTEN A. KERN

Electromagnetic actuators are the most frequently used actuator type within general automation-industry. Due to their simple manufacture and assembly they are the matter of choice. Additionally they do not necessarily need a permanent magnet, and their robustness against exterior influences is very high. They are used within coffee-machines, water pumps and for redirecting paper flow within office printers. But nevertheless their applicability for haptic devices, especially kinaesthetic devices, is limited. Their main fields of application are tactile systems. This can be reasoned by several special characteristics of the electromagnetic principle. Within this chapter the theoretical basics for electromagnetic actuators are given. Technical realizations are explained with examples, whereas first the general topology and later the specific designs are shown. The chapter closes with some examples of haptic applications of electromagnetic actuators.

9.3.1 Magnetic Energy

The source responsible for the movement of a magnetic drive is the magnetic energy. It is stored within the flux-conducting components of the drive. These components are given by the magnetic core (compare paragraphs 9.2.1.3 on page 203) and the air gap, such as all leakage fields - which are neglected for the following analysis. It is known from table 9.2 on page 204, that stored magnetic energy is given by the products of fluxes and magnetic voltages in each element of the magnetic circuit:

$$W_{\mathrm{mag}} = \sum_{n} H_n l_n \cdot B_n A_n \tag{9.37}$$

As every other system does, the magnetic circuit tries to minimize its inner energy[15]. Concentrating on electromagnetic actuators, the minimization of energy almost always refers to the reduction of the air gap's magnetic resistance R_{mG}. For this purpose two effects may be used, which can be found within electrostatic for electrical fields too (section 9.5):

- Electromagnetic longitudinal effect (fig. 9.24a) (also: reluctance effect)
- Electromagnetic transversal effect (fig. 9.24b)

The forces respectively torques generated with the individual effects are the derivations of the energy according to the corresponding direction,

[15] Minimizing potential energy is the basis for movements in all actuator principles. Actuators may therefore be characterized as *"assemblies aiming at the minimization of their inner energy"*.

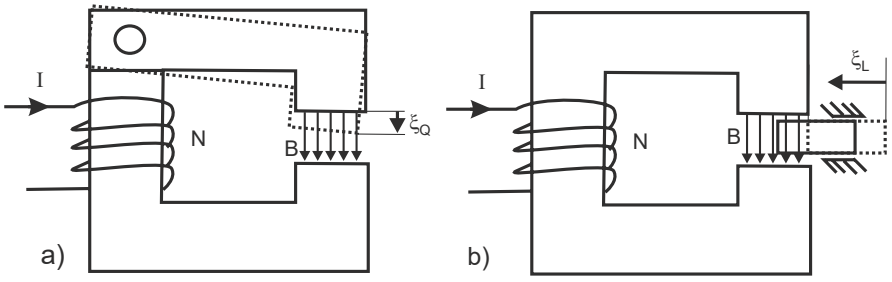


Fig. 9.24 Electromagnetic transversal- (a) and longitudinal-effect (b).

$$\mathbf{F}_\xi = \frac{dW_{\text{mag}}}{d\xi}, \tag{9.38}$$

being equal to a force in the direction of the change of the air gap

$$\mathbf{F}_\xi = -\frac{1}{2}\phi^2 \frac{dR_{\text{mG}}}{d\xi}. \tag{9.39}$$

Example: Transversal effect

The magnetic resistance of an arbitrary homogenous element of length l between two boundary surfaces (fig. 9.25a) with the surface A is calculated as

$$R_{\text{m}} = \frac{l}{\mu A}. \tag{9.40}$$

This gives the stored energy W_{mag} within the magnetic resistance:

$$W_{\text{mag}} = (BA)^2 \frac{l}{\mu A}. \tag{9.41}$$

The flux density B is dependent on the length of the material. Assuming that the magnetic core contains one material only the magnetomotive force Θ is calculated as

$$\Theta = \frac{B}{\mu l} = NI, \tag{9.42}$$

which gives

$$B = NI\frac{\mu}{l}. \tag{9.43}$$

This equation used to replace the flux density in equation 9.41, several variables canceled, finally results in the magnetic energy

$$W_{\text{mag}} = (NI)^2 A\mu \frac{1}{l}. \tag{9.44}$$

With the assumption about the magnetic energy concentrating within the air gap - which is identical to the assumption that the magnetic core does not have any relevant magnetic resistance - the approximation of the force for the transversal effect in the direction of l can be formulated

$$F_l = -\frac{1}{2}(NI)^2 A\mu \frac{1}{l^2}. \tag{9.45}$$

The force shows an anti-proportional quadratic coherence (fig. 9.25b) to the distance l. The closer the poles become, the higher the force attracting the poles increases.

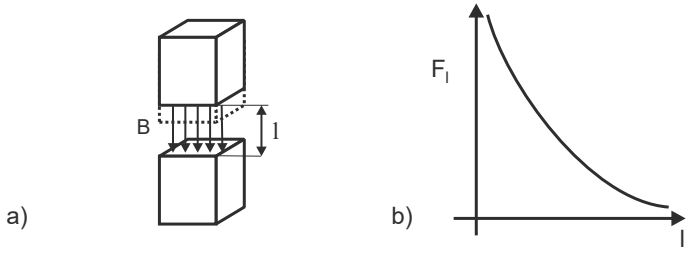


Fig. 9.25 Electromagnetic transversal effect in the air-gap (a) and with a qualitative force plot (b).

Example: Longitudinal effect, reluctance effect

The same calculation can be repeated for the longitudinal effect. Assuming that the surfaces A from equation 9.44 is rectangular and its edges' lengths are given by a and b, further assuming that a flux-conducting material is inserted into direction a, the forces in longitudinal direction can be calculated as

$$F_a = (NI)^2 b\mu \frac{1}{l}, \tag{9.46}$$

and in direction b as

$$F_b = (NI)^2 a\mu \frac{1}{l}. \tag{9.47}$$

The reluctance effect is - in contrast to the transversal effect - linear (fig. 9.26b). The force is dependent on the length of the mobbing material's edge only. Consequently the stored energy within the magnetic circuit is necessary for the design of an electromagnetic actuator. Above examples have the quality of rough estima-

tions. They are sufficient to evaluate the applicability of an actuation principle - no more, no less. The magnetic networks sufficient for a complete dimensioning should contain effects with magnetic leakage fields and the core's magnetic resistance. Therefore it is necessary to further deal with the design of magnetic circuits and their calculation.

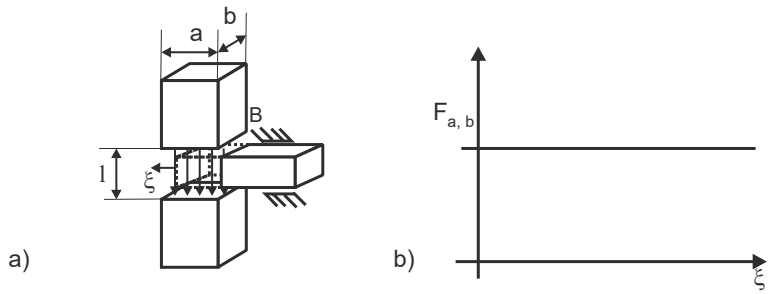


Fig. 9.26 Electromagnetic longitudinal effect in the air gap (a) and as qualitative force plot (b).

9.3.2 Design of Magnetic Circuits

The basic interdependencies for the design of magnetic circuits have already been discussed within paragraph 9.2.1.3 in the context of electrodynamic actuators. Taken from the approach of longitudinal- and transversal effect, several basic shapes (fig. 9.27) can be derived applicable to electromagnetic actuators. In contrast to electrodynamic actuators the geometrical design of the air gap within electromagnetic actuators is freer. There is no need to guide an electrical conductor within the air gap anymore. Beside the designs shown in figure 9.27 there are numerous other geometrical variants too. For example all shapes can be transferred into a rotational-symmetrical design around one axis. Additional windings and even permanent magnets can be added. There are just two limits to their design:

- A sufficient cross section of the flux conducting elements has to be guaranteed, to prevent the material from getting into saturation.
- A sufficient distance between flux-conducting elements has to be kept to prevent a magnetic leakage-field via the air.

9.3.2.1 Cross Section Surface Area - Rough Estimation

The calculation of the cross section surface area for dimensioning the magnetic core is simple. A common, easily available and within precision engineering and

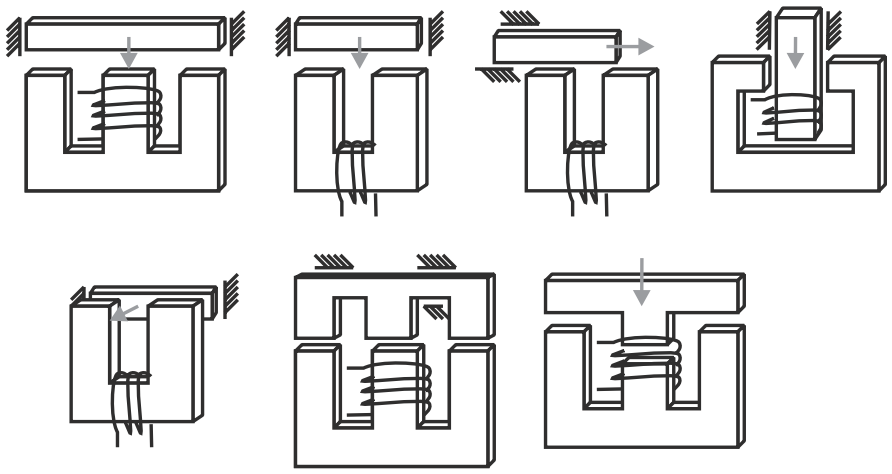


Fig. 9.27 Basic shapes of electromagnetic actuators.

prototype-design gladly used material is steal ST37. The B/H-characteristic curve with its saturation is given in figure 9.7 on page 210. For this example we choose a reasonable flux density of 1.2 T. This equals a field intensity of $H \approx 1000$ A/m. Within the air gap a flux density of 1 T should be achieved. The magnetic flux within the air gap is given as

$$\phi = A_{\mathrm{G}} B_{\mathrm{G}}. \tag{9.48}$$

As the magnetic flux is conducted completely via the magnetic core - neglecting leakage fields and other side bypasses - the relation

$$A_{\mathrm{Iron}} B_{\mathrm{Iron}} = A_{\mathrm{G}} B_{\mathrm{G}}, \tag{9.49}$$

is given, and consequently with the concrete values from above:

$$\frac{A_{\mathrm{Iron}}}{A_{\mathrm{G}}} = \frac{B_{\mathrm{G}}}{B_{\mathrm{Iron}}} = 0.833\,. \tag{9.50}$$

At its tightest point the magnetic core may have 83% of the cross section of the air gap. Whereas more surface of the cross-section results in lower field intensities, which should be aimed at if is geometrically possible. Please note that $A_{\mathrm{G}} \leq A_{\mathrm{Iron}}$ is with almost all technical realization, as the boundary surface of the magnetic core is always one pole of the air gap.

9.3.2.2 Magnetic Energy in the Magnetic Core and Air Gap

Within the preceding examples the assumption was made that the energy stored within the magnetic core is clearly less than the energy within the air gap. This assumption should be checked now for validity. Calculating the magnetic resistance of an arbitrary element

$$R_{\mathrm{m}} = \frac{l}{\mu A}. \tag{9.51}$$

the relation of two elements with identical length and cross-section scales via the magnetic resistance with the permeability μ:

$$\frac{R_{\mathrm{m1}}}{R_{\mathrm{m2}}} = \frac{\mu_2}{\mu_1}. \tag{9.52}$$

The permeability $\mu_r = \frac{B}{H\mu_0}$ is given by the relation between flux density vs. field strength relatively to the magnetic constant. It is nonlinear (fig. 9.28) for all typical flux-conducting materials within the flux-density areas relevant for actuator design in between 0.5 and 2 T. It is identical to the inverse gradient of the curves given in diagram 9.7. The maximum permeability values are given in tables frequently, but refer to field strengths within the material only. They range from 6,000 for pure iron over 10,000 for nickel alloys up to 150,000 for special soft-magnetic materials.

> Mechanical processing of flux-conducting materials and the resulting thermal changes within its microstructure will result in a considerable degradation of its magnetic properties. This change can be restored by an annealing process.

Generally speaking however even outside an optimum value for flux density the stored energy within typical materials is always several orders of magnitudes below the energy within the air gap. This legitimates to neglect this energy portion for the rough estimations in actuator design, but does show too that there is potential within the optimization of electromagnetic actuators. This potential can be used by the application of FEM-software, which is typically available as module for CAD software [16].

9.3.2.3 Permanent Magnets in Electromagnetic Actuators

Permanent magnets do not differ significantly in their properties from coils conducting a DC current. They generate a polarized field, which - in combination with another field - provide attraction or repulsion. For calculating combined magnetic circuits a first approach can be taken by substituting the sources within the magnetic equivalent circuit (neglecting saturation effects). The calculation is analog to

[16] or as free software, e.g. the tool “FEMM” from DAVID MEEKER

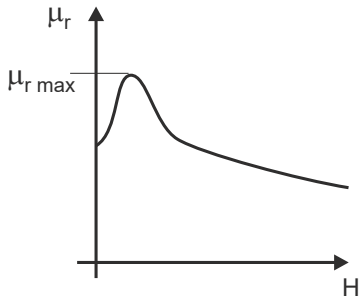


Fig. 9.28 Qualitative change of the permeability for common flux-conducting materials.

the methods read in the chapter about electrodynamic actuators (paragraph 9.2.1.3). A permanent magnet within a circuit either allows

- the realisation of a state held without any current applied
- or switching between two states with just one winding powered

A good example for a currentless held state [125] shows the calculation of a polarized magnetic clamp (fig. 9.29). With non-active winding the flux is guided through the upper anchor and held securely. With active coil the magnetic flux via the upper anchor is compensated. The magnetic bypass above the coil prevents the permanent magnet from being depolarized by a counter field beyond the kink in the B/H-curve.

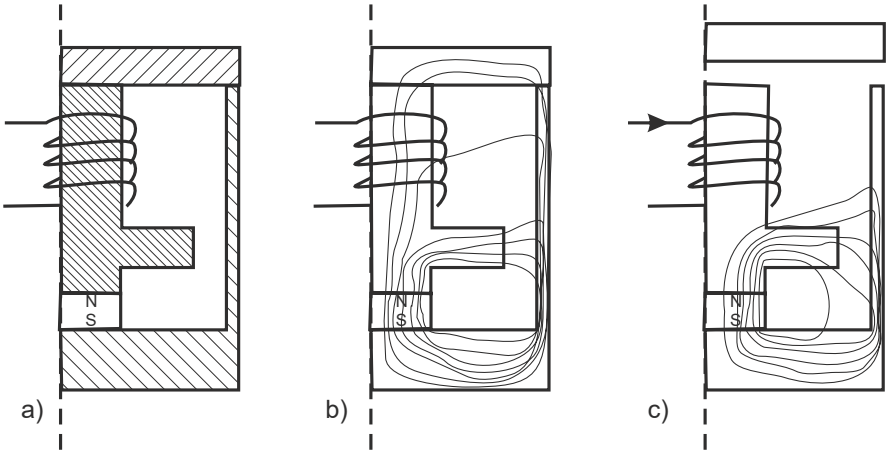


Fig. 9.29 Permanent magnet in the magnetic circuit in shape (a), field-lines with inactive coil b) and field-lines with active coil (c), releasing the anchor.

9.3.3 Examples for Electromagnetic Actuators

Electromagnetic actuators are available in many variants. The following section gives typical designs for each principle and shows the corresponding commercial products. The knowledge about the designs will help to understand the freedom in the design of electromagnetic circuits more broadly.

9.3.3.1 Claw-pole Stepper Motor

The electromagnetic claw-pole stepper motor (fig. 9.30) is one of the most frequently used rotatory actuation principle. These actuators are made of two stamped metal sheets (1,2) with the poles - the so called claws - bended by 90° to the interior of the motor. The metal sheets are the magnetic core for conducting the flux of one coil each (3). The permanent-magnet rotor (4) with a pole subdivision equalizing the claw pattern orientates to the claws in currentless state. In stepper mode the coils are powered subsequently, resulting in a combined attraction and repulsion of the rotor. The control of the coils' currents may happen either by simply switching them or by a microstep mode with different current levels being interpolated between discrete steps. Latter generates stable states for the rotor not only in the positions of the claws but also in between.
Claw-pole stepper motors are available with varying number of poles, different numbers of phases and varying loads to be driven. As a result of the permanent-magnet they show a large holding torque with respect to their size. The frequency of single steps may reach up to 1 kHz for fast movements. By counting the steps the position of the rotor can be detected. Step-losses - the fact that no mechanical step happens after a control signal - are not very likely with a carefully designed power chain. Claw-pole stepper motors are the working horses of the electrical automation technology.

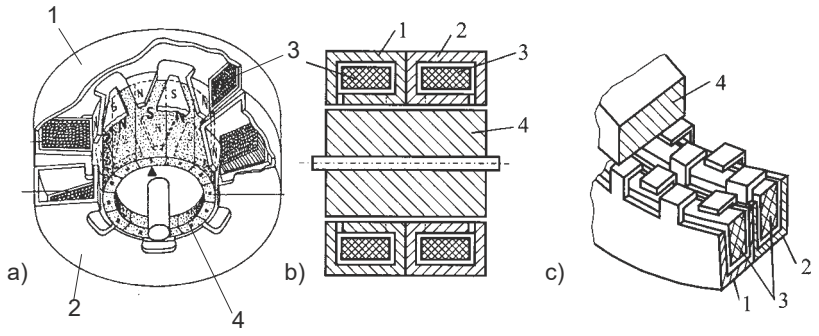


Fig. 9.30 Two-phase stepper motor made of stamped metal sheets and with a permanent magnet rotor in a 3D-sketch (a), cross-section (b), and with details of the claws-poles (c) [218].

9.3.3.2 Reluctance Drives

The rotatory reluctance drives (fig. 9.31) are based on the electromagnetic longitudinal effect. By clever switching of the windings (2) it is possible to keep the rotor (3) in a continuous movement with just minimal torque ripples. To make this possible the rotor has to have fewer poles than the stator. The rotor's pole-angle β_r is larger than the pole-angle β_s. Reluctance drives can be used as stepper motors by the integration of permanent magnets too. Generally speaking it excels by high robustness of the components and a high efficiency factor with - for electromagnetic drives - comparably little torque ripples.

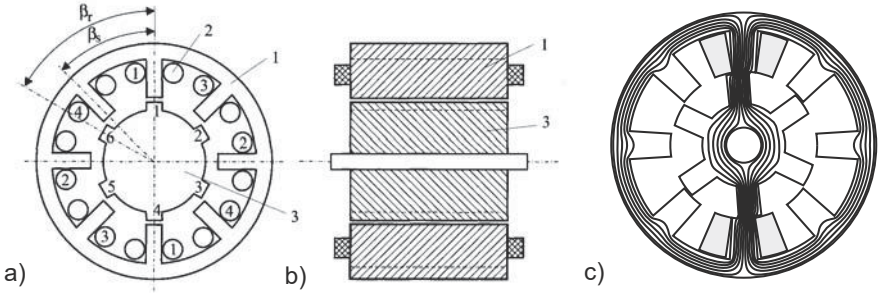


Fig. 9.31 Switched reluctance drive with pole- and coil-layout (a), in cross section (b), and with flux-lines of the magnetic excitation (c) [218].

9.3.3.3 Electromagnetic Brakes

Electromagnetic brakes (fig. 9.32) are based on the transversal effect. They make use of the high force increase at electromagnetic attraction to generate friction on a spinning disk (1). For this purpose usually rotational-symmetrically flux conducting magnetic cores (2) are combined with embedded coils (3). The frontal area of the magnetic core and braking disk (1) itself is coated with a special layer to reduce abrasion and influence positively the reproducibility of the generated torque. The current/torque characteristic of electromagnetic brakes is strongly nonlinear. At the one hand this is the result of the quadratic proportion between force and current of the electromagnetic transversal effect, but on the other hand this is also a result of its friction pairing. Nevertheless they are used in haptic devices for the simulation of "hard contacts" and stop positions. A broad application for haptic devices is nevertheless not visible. This is likely a result of their limits in reproducibility of the generated torque, the resulting complex control of the current and the fact, that they can only be used as a break (passive) and not for active actuation.

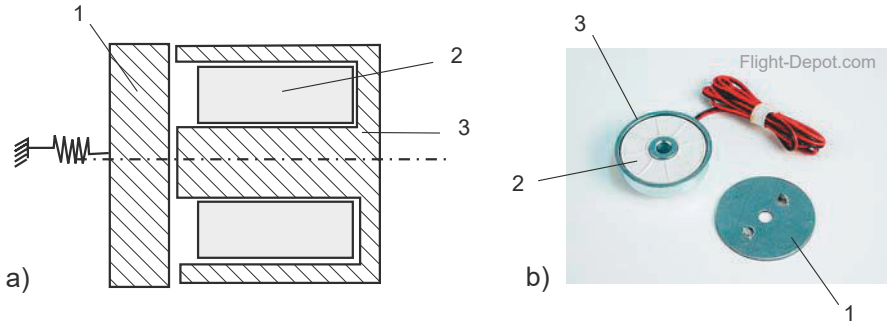


Fig. 9.32 Electromagnetic brake in cross section (a) and as technical realization for an airplane model (b).

9.3.3.4 Plunger-type Magnet

Electromagnetic plunger-type magnets (fig. 9.33) are frequently based on the electromagnetic transversal effect. Their main uses are switching and control applications, requiring the actuation in specific states. With a magnetic core frequently made of bended iron steel sheets (2) a coil-induced (3) flux is guided to a central anchor, which itself is attracted by a yoke (4). The geometry of the yoke influences significantly the characteristic curve of the plunger-type magnet. By varying its geometry a linearization of the force-position curve is possible within certain limits. Even strongly nonlinear pulling-force characteristics can be achieved by such a modification. Plunger-type magnets are available with additionally magnets and with more coils. In these more complex designs they provide mono- and bi-stable switching properties. By variation of the wires diameter and the number of turns they can be adapted easily to any power level.

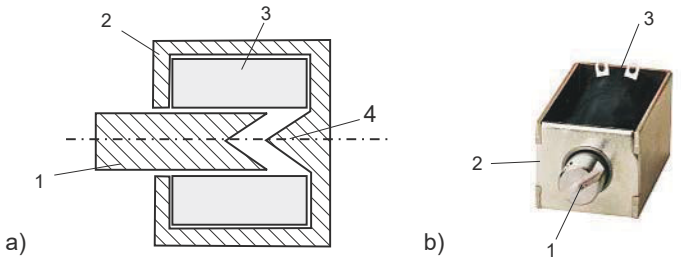


Fig. 9.33 Plunger type magnet (a) with altered force-position curve (4), and realization as pulling-anchor (b) with metal-sheet-made magnetic circuit (2).

9.3.4 Magnetic Actuators in Haptic Devices

For haptic applications electromagnetic actuators are mainly used within tactile devices. Nevertheless admittance controlled devices can be found providing impressive output forces of high quality even by the use of stepper-motors. Beside the commercial systems such as "HapticMaster" of Moog (section 5.1.4) especially an idea of LAWRENCE attracted attention within the past few years.

9.3.4.1 Spring-Tendon Actuator

In [157] LAWRENCE describes an inexpensive actuator for kinaesthetic haptic systems (fig. 9.34) based on an electromagnetic stepper motor coupled via a tendon to a pen and with a spring mechanically connected in parallel. Analog to other haptic devices the pen is the interface to the user. Between pen and tendon and spring there is a bending body with DMS as a force sensor. To additionally minimize the torque ripples of the stepper drive resulting from the latching of the poles, a high resolution external encoder has been attached to the motor, and a torque/angle curve was measured. A mathematical spline fit of this curve was used for the actuator's control to compensate the torque oscillations. Beside this compensation the closed-loop control of the actuator via the force sensor near the pen includes a compensation of frictional effects too. The result of all these efforts is a force source, providing a force transmission with little noise and high stiffness up to 75 kN/m with movements of limited dynamics.

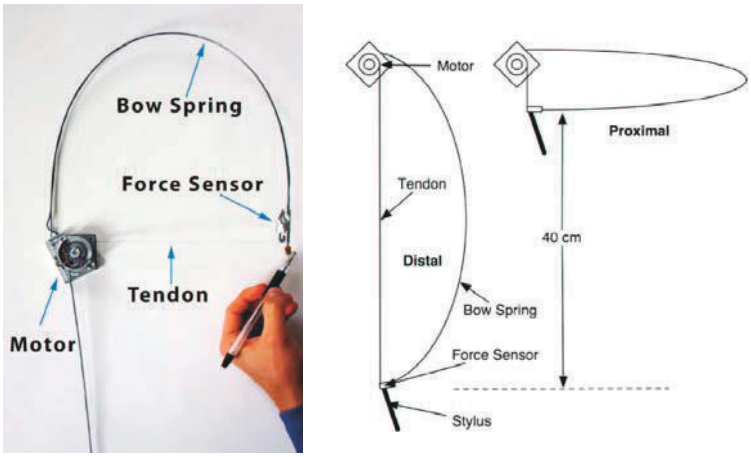


Fig. 9.34 Electromagnetic stepper motor with spring, actuated in closed-loop admittance control mode [157].

9.3.4.2 Electromagnetic Pin Array

The usage of electromagnetic actuators for the control of single pins in array design is very frequent. The earliest usages for haptic applications go back to printer-heads of dot matrix printers being used in the 80th an earl 90th of the last century. Modern designs are a lot more specific to haptics and make use of manufacturing technologies available from microtechnology. In [45] an actuator array (fig. 9.35) is shown made of coils with 430 windings each and a 0.4 mm wide iron core. Above of it a magnet is embedded in a polymer-layer, being attracted by the flux induced into the core. With such an actuator and a diameter of 2 mm a maximum force of up to 100 mN is possible. A position measure is realized via inductive feedback.

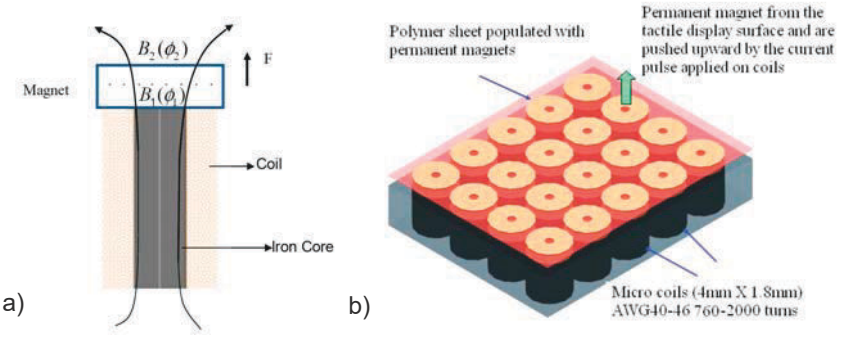


Fig. 9.35 Electromagnetic monostable actuator with permanent magnet: principle sketch (a) and actuator design (b) [45].

9.3.4.3 Electromagnetic Plunger-type Magnet for the Tactile Transmission of Speech Information

One fundamental motivation for the design of haptic devices is the partially substitution of lost senses. Especially methods to communicate information from the sense of sight or hearing by the aid of tactile devices have some tradition. BLUME designed and tested an electromagnetic plunger-type magnet according to the reluctance effect 1986 at the University of Technology, Darmstadt. Such actuators were attached to the forearm and stimulated up to eight points by mechanical oscillations encoded from speech signals. The actuator (fig. 9.36) was made of two symmetrical plunger-type magnets (on the horizontal axis) acting upon a flux conducting element integrated into the plunger. The whole anchor takes a position symmetrically within the actuator due to the integrated permanent magnet. In this symmetrical position both magnetic circuits conduct a magnetic flux resulting in identical reluctance forces. In active mode either the flux in the upper or the lower magnetic circuit

is amplified depending on the direction of current flow. The reluctance forces on the amplified side pull the anchor in a current proportional position working against the magnetic pretension from the permanent magnets, the mechanic pretension from the springs and the load of the actuator. The plunger is displaced in the direction of the weakened magnetic field. At a diameter of 20 mm this actuator covers a dynamic range of 500 Hz at a efficiency factor of 50%. The forces lie in the range of ≈4 N per ampere.

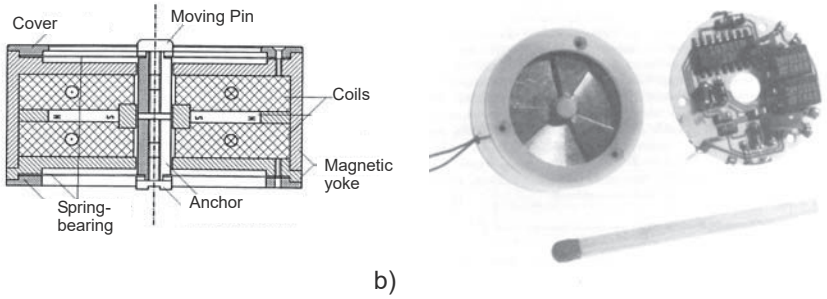


Fig. 9.36 Electromagnetic actuator according to the reluctance principle in a "counteractive plunger type" design with permanent magnet: cross-section (a) and design incl. driver-electronics (b) [25].

9.3.5 Conclusion about the Design of Magnetic Actuators

Electromagnetic actuators are - identical to electrodynamic systems - mainly force sources. In rotary drives especially the reluctance effect is used to generate a continuous movement. With linear drives mainly plunger-type magnets are used based on the nonlinear transversal effect. Whereas there are exception to both showing some surprising properties (paragraph 9.3.4.3). The translational systems are usually used to actuate as either bistable switches between two discrete states or monostable against a spring (plunger-type, break, and valve). There are applications within haptics based on either of both effects. Whereas reluctance based actuators can be found equally often within kinaesthetic applications as drives and in an admittance controlled application in tactile systems as vibrating motor, switching actuators are almost exclusively found in tactile devices with single pins or pin arrays. In contrast to the highly dynamic electrodynamic drives, electromagnetic actuators excel in less dynamic applications with higher requirements on torque and self-holding. During switching in between two states however the acceleration and deceleration at the mechanical stop are a highly dynamic but almost uncontrollable action. The dynamic design of switching actions were not subject to this chapter, but are usually

based on modeling a nonlinear force source of the electromagnet and assuming the moving parts as concentrated elements of masses, springs and dampers. Due to their relatively high masses within the moving parts, the hard to control nonlinearities of fluxes and forces, and the low efficiency factor of the transversal effect in many designs, electromagnetic actuators occupy niches within haptic applications only. However in those niches there is no way around their usage. If an appropriate area has been found they excel by an extremely high efficiency factor for the specific design and a big robustness against exterior influences.

9.4 Piezoelectric Actuators

STEPHANIE SINDLINGER, MARC MATYSEK

Next to the very frequently found electrodynamic actuators, the past few years piezoelectric actuators were used for a number of device designs. Especially their dynamic properties in resonance mode allow an application for haptics, which is very different from the common positioning application they are used for. As variable impedance a wide spectrum of stiffnesses can be realized. The following chapter gives the calculation basics for the design of piezoelectric actuators. It describes the design variants and their application in haptic systems. Beside specific designs for tactile and kinesthetic devices approaches for the control of the actuators and tools for their dimensioning are presented.

9.4.1 The Piezoelectric Effect

The piezoelectric effect was discovered by JACQUES and PIERRE CURIE first. The term is derived from the Greek word “piedein - piezo” = “to press” [116].

Figure 9.37 shows a scheme of a quartz crystal (chemical: SiO2). With force acting upon the crystal mechanical displacements of the charge-centers can be observed within the structure, resulting in microscopic dipoles within its elementary cells. All microscopic dipoles sum up to a macroscopic measurable voltage. This effect is called “reciprocal piezoelectric effect”. It can be reversed to the “direct piezoelectric effect”. If a voltage is applied on a piezoelectric material a mechanical deformation happens along the crystal’s orientation, which is proportional to the field strength in the material [7, 8, 225].

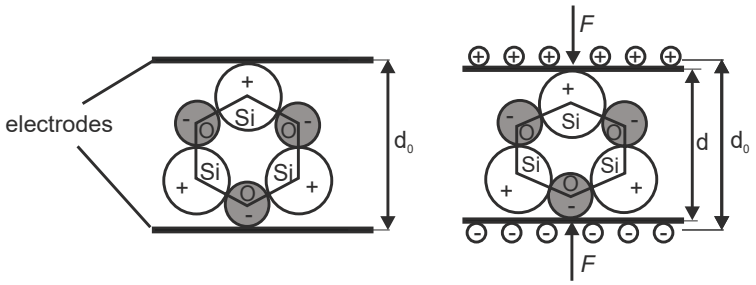


Fig. 9.37 Crystal structure of quarz in initial state and under pressure [225]

Piezoelectric materials are anisotropic - direction dependent - in their properties. Consequently the effect depends on the direction of the electrical field applied, and on the angle between the direction of the intended movement and the plane of polarization. For the description of these anisotropic properties the directions are labeled with indices. The index is defined by a Cartesian space with the axes being numbered with 1, 2 and 3. The plane of polarization of the piezoelectric material is typically orientated on direction 3. The shear at the axes is labeled with indices 4, 5 and 6.

Among all possible combinations, there are three major effects (fig. 9.38), commonly used for piezoelectric applications: longitudinal- , transversal- and shear-effect.

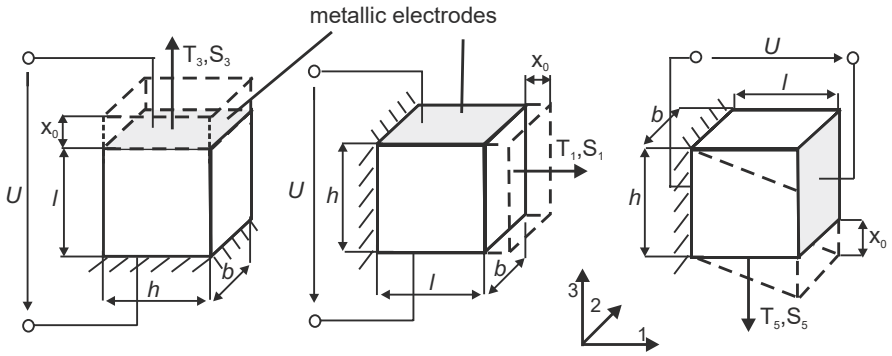


Fig. 9.38 Effects during applied voltage: longitudinal effect (left), transversal effect (center), shear effect (right)[225]

The *longitudinal effect* acts in the same direction as the applied field and the corresponding field strength E_3. As a consequence the resulting mechanical tensions T_3 and strains S_3 lie within plane 3 too. With the *transversal effect* mechanical actions show normal to the electrical field. As a result from a voltage U_3 with the electrical field strength E_3 the mechanical tensions T_1 and strains S_1 appear. The *shear-effect* happens with the electrical voltage U applied along plane 1 orthogonal to the polarization plane. The resulting mechanical tensions appears tangential to the polarization - in the direction of shear - and are labeled with the directional index 5.

9.4.1.1 Basic Piezoelectric Equations

The piezoelectric effect can be described most easily by state equations:

$$P = e \cdot T \tag{9.53}$$

and

$$S = d \cdot E \tag{9.54}$$

with

P = direction of polarization (in C/m^2)
S = deformation (non-dimensional)
E = electrical field strength (in V/m)
T = mechanical tension (in N/m^2)

The piezoelectric coefficients are

- the piezoelectric coefficient of tension (also: coefficient of force) e (reaction of the mechanical tension on the electrical field)

$$e_{ij,k} = \frac{\partial T_{ij}}{\partial E_k} \partial \tag{9.55}$$

- and the piezoelectric coefficient of strain (also: coefficient of charge) d (reaction of the strain on the electrical field)

$$d_{ij,k} = \frac{\partial \varepsilon_{ij}}{\partial E_k} \partial \tag{9.56}$$

The correlation of both piezoelectric coefficients is defined by the elastic constants C_{ijlm}

$$e_{ij,k} = \sum_{lm} \left(C_{ijlm} \cdot d_{lm,k} \right) \tag{9.57}$$

Usually the tensors shown in the equation above are noted as matrix (VOIGT'sche notation) . In this format, matrices result of six components identical to the defined axes. The matrix shown below describes the concatenation of the dielectrical displacement D, the mechanical strain S, the mechanical tension T, and the electrical field strength E.

	T_1	T_2	T_3	T_4	T_5	T_6	E_1	E_1	E_3
D_1	0	0	0	0	d_{15}	0	ε_{11}	0	0
D_2	0	0	0	d_{15}	0	0	0	ε_{11}	0
D_3	d_{31}	d_{31}	d_{33}	0	0	0	0	0	ε_{11}
S_1	s_{11}	s_{12}	s_{13}	0	0	0	0	0	d_{31}
S_2	s_{12}	s_{11}	s_{13}	0	0	0	0	0	d_{31}
S_3	s_{13}	s_{13}	s_{33}	0	0	0	0	0	d_{33}
S_4	0	0	0	s_{44}	0	0	0	d_{15}	0
S_5	0	0	0	0	s_{44}	0	d_{15}	0	0
S_6	0	0	0	0	0	$2(s_{11} - s_{12})$	0	0	0

This matrix can be simplified for the specific cases of a longitudinal and a transversal actuator. For a longitudinal actuator with electrical contact in direction 3 the following equations are the result .

$$D_3 = \varepsilon_{33}^T E_3 + d_{31} T_1 \tag{9.58}$$

$$S_3 = d_{31} E_3 + s_{11}^E T_1 . \tag{9.59}$$

Accordingly for a transversal actuator the correlation

$$D_3 = \varepsilon_{33}^T E_3 + d_{33} T_3 \tag{9.60}$$

$$S_3 = d_{33} E_3 + s_{33}^E T_3 \tag{9.61}$$

with

D_3	=	dielectric displacement in C/m^2	D=0: open-ended
E_3	=	field-strength in V/m	E=0: short-cut
S_1, S_3	=	$\Delta L/L$ = strains, dimensionless	S=0: mech. short-cut
T_1, T_3	=	mechanical tensions N/m^2	T=0: idle mode
ε_{33}^T	=	relative dielectricity constant at mechanical tension = 0	
d_{31}, d_{33}	=	piezoelectric charge constant in C/N	
s_{11}^E, s_{33}^E	=	elasticity constant at field strength = 0	

becomes valid.

Therefore the calculation of piezoelectric coefficients simplifies into some handy equations:

The charge constant d can be calculated for the electrical short-circuit- which is $E = 0$- to

$$d_{E=0} = \frac{D}{T} \tag{9.62}$$

and for the mechanical idle situation - which is $T = 0$ - to

$$d_{T=0} = \frac{S}{E} . \tag{9.63}$$

The piezoelectric tension constant is defined as

$$g = \frac{d}{\varepsilon^T} . \tag{9.64}$$

The coupling factor k is given by equation 9.65. It is a quantity for the energy transformation and consequently for the strength of the piezoelectric effect. It is used for comparison among different piezoelectric materials. However note that it is not identical to the efficiency factor, as it does not include any energy losses.

$$k = \frac{converted\ energy}{absorbed\ energy}. \tag{9.65}$$

A complete description of the piezoelectric effect, a continuative mathematical discussion, and values for piezoelectric constants can be found in literature, such as [29, 104, 158].

9.4.1.2 Piezoelectric Materials

Till 1944 the piezoelectric effect was observed with monocrystals only. These were quartz, turmalin, lithiumniobat, potassium- and ammonium-hydrogen-phosphat (KDP, ADP), and potassium sodium tartrate [7]. With all these materials the direction of the spontaneous polarisation is given by the direction of the crystal lattice [116]. DThe most frequently used material was quartz.

The development of polarization methods made it possible to retrospectively polarize ceramics by the application of a constant exterior electrical field in 1946. By this approach “piezoelectric ceramics” (also “piezoceramics”) were invented. By this development of polycrystalline materials with piezoelectric properties the whole group of piezoelectric materials got an increased attention and technical significance. Today the most frequently used materials are e.g. barium titanate (MaTiO3) or lead zirconate titanate (PZT) [7]. C 82 is a piezoelectric ceramic suitable for actuator design due to its high *k-factor*. However as all piezoelectric ceramic materials it shows reduced long term stability compared to quartz. Additionally it has a pyroelectric effect which is a charge increase due to temperature changes of the material [158]. Since the 60th the semi-crystalline synthetic material polyvinylidene fluoride (PVDF) is known. Compared to the materials mentioned before, PVDF excels by its high elasticity and reduced thickness (6 μm to 9 μm).

Table 9.4 shows different piezoelectric materials with their specific values.

Looking at these values PZT is most suitable for actuator design due to its high coupling factor with large piezoelectric charge modulus and still a high Curie temperature . The Curie temperature represents the temperature at which the piezoelectric properties from the corresponding material are lost permanently. The value of the curie temperature depends on the material (table 9.4).

9.4.2 Designs and Properties of Piezoelectric Actuators

Actuators using the piezoelectric effect are members of the group of solid actuators (also: solid-state actuators). The transformation from electrical into mechanical energy happens without any moving parts, resulting in a very fast reaction time and high dynamics compared to other actuation principles. Additionally piezoelectric actuators have a high durability. The thickness changes are smaller - compared to other actuation principles. Although the generated forces are much higher.

Table 9.4 Selection of piezoelectric materials with characteristic values[158].

Constant	Unit	Quartz	PZT-4	PZT-5a	C 82	PVDF
d_{33}	$\frac{10^{-12}}{m/V}$	2,3	289	374	540	-27
d_{31}		-2,3	-123	-171	-260	20
e_{33}	$\frac{A \cdot s}{m^2}$	0,181	15,1	15,8	28,1	108
e_{31}		-0,181	-5,2	-5,4	-15,4	-
s_{33}^E	$\frac{10^{-12}}{m^2/N}$	12,78	15,4	18,8	19,2	-
s_{11}^E		12,78	12,3	16,4	16,9	-
c_{33}^E	$\frac{10^{10}}{N/m^2}$	7,83	6,5	5,3	5,2	-
c_{11}^E		7,83	8,1	6,1	5,9	-
$\frac{\varepsilon_{33}^T}{\varepsilon_0}$; $\frac{\varepsilon_{33}^S}{\varepsilon_0}$	-	4,68; 4,68	1300; 635	1730; 960	3400; -	12 ;12
$\frac{\varepsilon_{11}^T}{\varepsilon_0}$; $\frac{\varepsilon_{11}^S}{\varepsilon_0}$	-	4,52; 4,41	1475; 730	1700; 830	3100; -	-
k_{33}	-	0,1	0,7	0,71	0,72	0,20
k_{31}	-	-	0,33	0,34	0,36	0,15
ϑ_{Curie}	$\frac{1}{^\circ C}$	575	328	365	190	80
ρ	$\frac{kg}{m^{-3}}$	2660	7500	7500	7400	1790

9.4.2.1 Basic Piezoelectric Actuator Designs

Depending on the application different designs may be used. One may require a large displacement; another one may require self-locking or high stiffnesses. The most frequently used actuator types are bending actuators and actuator staples. A schematic sketch of each design is given in figure 9.39 (a) and (c).

Stacked actuator are based on the longitudinal piezoelectric effect. For this purpose several ceramic layers of opposite polarity are stapled above each other. In between each layer contact electrodes are located for electrical control. A staple is able to generate high static forces up to several 10 kN. The achievable displacement of 200 μm is low compared to other piezoelectric designs. By the use of levers 9.39 (b) the displacement can be significantly increased (s. figure 9.39 (b)). Voltages of several 100 V are necessary to drive a piezoelectric actuation staple.

Bending actuators are based on the transversal piezoelectric effect. Designed according to the so called bimorph principle - with two active layers - they are used in applications requiring large displacements. The transversal effect is characterized by comparably low controlling voltages [7, 116]. These electrical properties and the large displacements can be achieved by very thin ceramic layers in the direction of the electrical fields, and an appropriate geometrical design. Other geometrical designs using the transversal effect are tubular actuators, film actuators, or bending discs 9.39 (d) to (f). Due to their geometry they equal staple actuators in their mechanical and electrical characteristics. The achievable displacements of 50 μm are

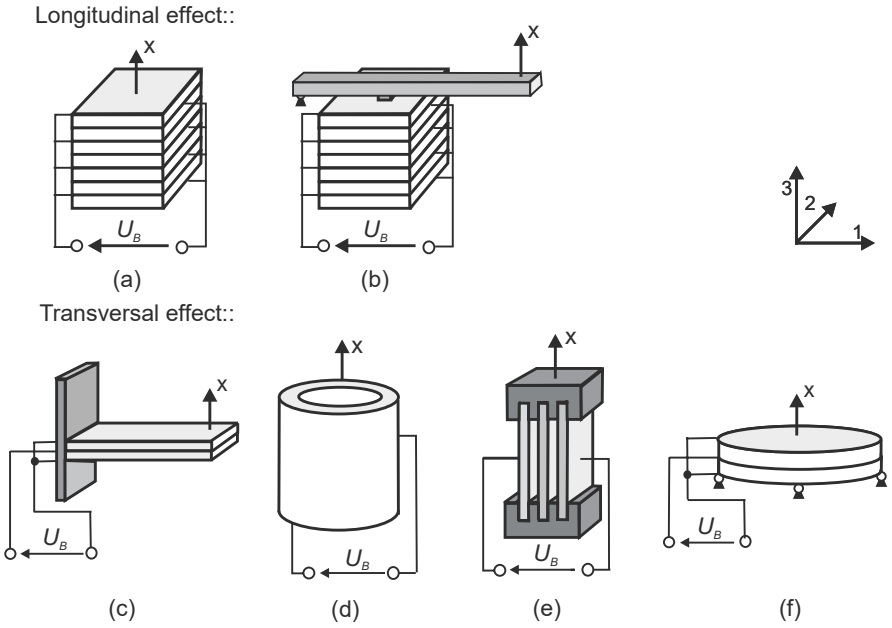


Fig. 9.39 Piezoelectric transducers separated by longitudinal and transversal effect: longitudinal effect: (a) stack, (b) stack with lever transformation, change of length: $x = d_{33} \cdot U_B$ transversal effect: (c) bending actuator, (d) cone, (e) band (f) bending disk, change of length: $x = -d_{31} \cdot U_B$ [116]
.

comparably low, whereas the achievable forces excel bending actuators at several orders of magnitude..

The use of the *shear effect* is uncommon in actuator design. This is somewhat surprising as it shows charge modulus and coupling factor which is twice as much as the transversal effect. Additionally it is possible to increase the elongation x_0 in idle mode (displacement without any load) by the optimization of the length to thickness (l/h) ratio. However the clamping force F_k of the actuator is not influenced by these parameters.

Table 9.5 summarizes the properties of different geometrical designs. Typical displacements, actuator forces and control voltages are shown.

Table 9.5 Properties of typical piezoelectric actuator designs [116].

Standard designs	stack	stack with lever	bending actuator	tape actuator	tubus	bending discs
actuator displacements	20...200 μm	$\leq$ 1.000 μm	$\leq$ 1.000 μm	$\leq$ 50 μm	$\leq$ 50 μm	$\leq$ 500 μm
actuating forces	$\leq$ 30.000 N	$\leq$ 3.500 N	$\leq$ 5 N	$\leq$ 1000 N	$\leq$ 1000 N	$\leq$ 40 N
control voltages	60...200 V 200...500 V 500...1000 V	60...200 V 200...500 V 500...1000 V	10...400 V	60...500 V	120...1000 V	10...500 V

9.4.2.2 Selection of Special Designs for Piezoelectric Actuators

Beside the standard designs shown above several variations of other geometrical designs exist. In this section, examples of an ultrasonic drive with resonator, oscillatory/traveling waves actuators and piezoelectric stepper motors are discussed. Ultrasonic actuators are differentiated according to resonators with bar like geometry and rotatory ring geometry.

Ultrasonic Actuators with Bar-shaped Resonator

One special variant of piezoelectric ultrasonic actuators is made of a piezoelectric actuator staple and a mechanical resonator [116, 260, 261, 263, 280]. The resonator is part of the stator of the drive, and its oscillating end actuates the rotor. As the movement of the actuator is not continuous and resembles represents a standing wave, the design is also known as “standing-wave actuator”. ,

The oscillating movement of the stator is transferred on the rotor and makes it move. Although the amplitudes are in the area of several μm, rotational speeds of 100 turns / minute are easily possible:

The amplitude at the end of the resonator is $4\,\mu$m. With an oscillation of $25\,kHz$ the tip moves

$$v = 4\,\mu m \cdot 25\,kHz = 100\,m/sec. \tag{9.66}$$

Basically two different designs of bar like resonators are available: mono-modal and bi-modal actuators. Both designs are schematically sketched in figure 9.40.

With mono-modal actuators only one oscillatory resonance mode is builds up. It is actuated in one frequency, resulting in a longitudal movement of the bar’s tip in axial direction. The tip itself is beveled. The contact point between bar and rotor is placed asymmetrically to the rotary axis. The one-dimensional oscillation of the axis results in a push-movement against the rotor. To change the direction of movement the position of the bar has to change (or a second bar has to be available) pushing the

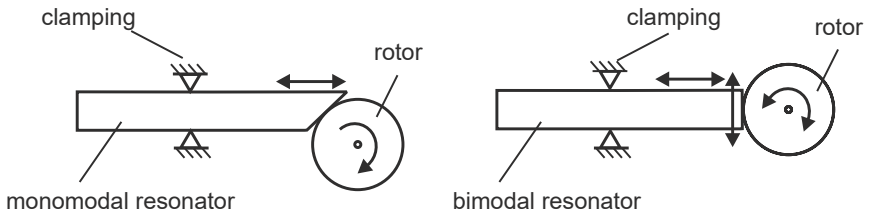


Fig. 9.40 Design of the monomodal (left) and bimodal (right) piezoelecric actuator with a bar-shaped resonator [116].
.

rotor below the rotatory axis. With bi-modal actuators the resonator is used in two of its resonance modes for oscillation: one longitudinal mode and one bending mode. The control is done with two frequencies being summed up in one signal. With the frequencies being chosen according to one frequency resembling an integer multiple of the other one, a closed rotary motion results (LISSAJOUS-figure). By changing the phase lag between both frequencies a rotation in both direction is possible to achieve. Additionally the phase lag is able to alter the rotational speed of the motor. Figure 9.41 visualizes this correlation.

The geometrical shape of the tip's movement is defined by the phase Θ. With a phase angle $\Theta = 0$a optimum contact between tip and rotor is achieved for anti-clockwise movement. The rotational speed is at its maximum. By increasing Θ, the movement of the tip and accordingly the contact time and shape between resonator and rotor change. An example is given for $\Theta = \frac{\pi}{4}$. With a phase angle of $\Theta = \frac{\pi}{2}$ no driving torque is left. This phase can be used to prevent a self-locking of the motor in an idle mode. It is somewhat like an active unblocking state, wherein the friction between rotor and resonator is canceled by the oscillating movement of the resonator's tip. Another increase of the phase lag to $\Theta = \frac{3}{4}\pi$results in a change of direction for the motor. At $\Theta = \pi$the clockwise movement direction is done with an optimum in contact states and a maximum torque being transmitted.

The movement of the rotor tip according to figure 9.41 can be described by the following equation:

$$x = x_0 sin(\omega_1 t),\ y = y_0 \sin(\omega_2 t + \Theta). \tag{9.67}$$

Whereas f_1 dis the frequency of the longitudinal movement and $f_2 = 2 \cdot f_1$ the frequency of the bending mode.

If the bending mode is not designed to be an integer multiple of the transversal mode, the frequencies will drift during operation. In such cases a defined phase has to be identified between both control frequencies to keep a defined contact between tip and rotor. Considerations dealing with the optimization of the contact properties are given e.g. in [62, 80, 180, 248, 297].

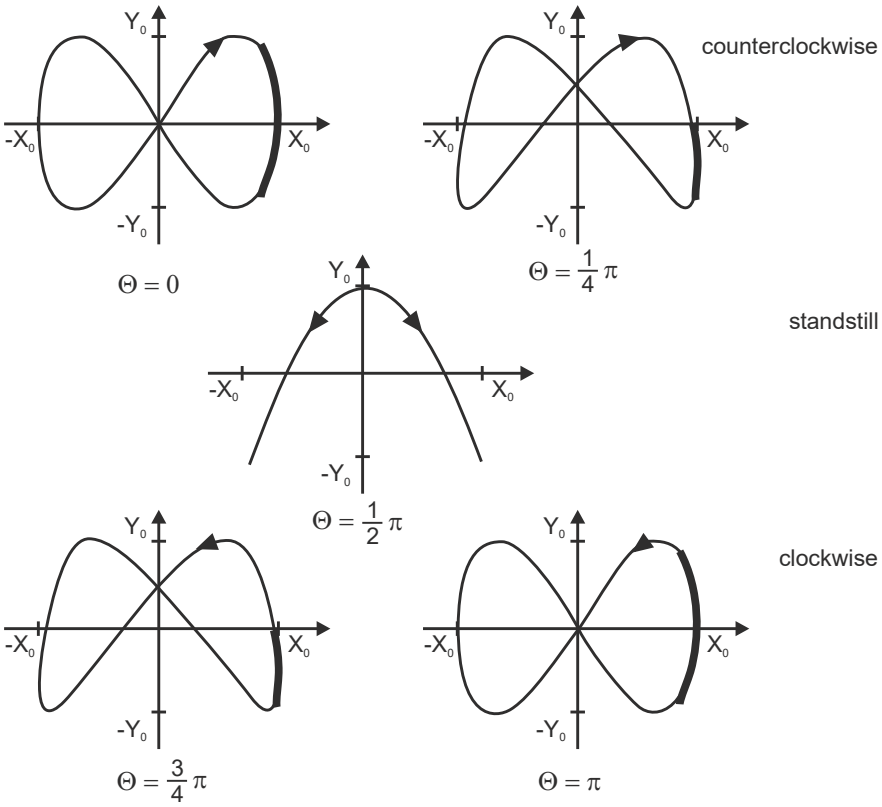


Fig. 9.41 Motion of the tip of the resonator of the bimodal resonator with varying phase shift. The contact area is marked. [116].
.

The movements' amplitudes of piezoelectric actuators with bar-like resonator lie in the area of several μm. To increase it the geometrical shape of the resonator has to be changed. As the resonator is also called "horn", as it fulfills the same functionality as the horn of any wind instrument - it changes the amplitude of the input oscillation. There are three basic shapes common for horns: the stepwise narrowing horn, the exponential narrowing horn and the linear narrowing horn. The geometrical design of horns is challenging and goes widely beyond the scope of this book. Additional information on the shapes of horns can be taken from [16, 260, 261, 263]. Calculation hints and optimizing criterions for the design of resonators for piezoelectric actuators are discussed in [48, 228, 262, 280].

A commercially available piezoelectric actuator with mono-modal resonator and still able to move in two directions is sold by the company Elliptec [168, 268].

The shape of this resonator and the placement of the piezoelectric elements are - compared to the basic design - varied 9.42.

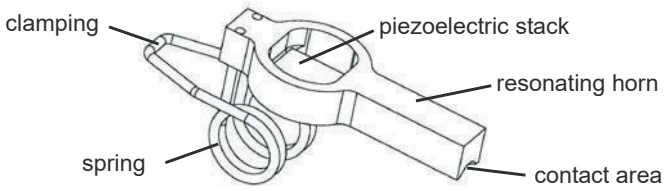


Fig. 9.42 Drawing of the piezoelectric ultrasonic actuator by Elliptec [168].
.

The motor combines a piezoelectric staple with resonator, whose tip is oscillating elliptically. There is a groove included in the tip providing a contact surface and guide between driven element and the resonator. The actuator is pressed by an elastic element, e.g. a torsion- or flat-spring, at a defined angle against the driven element. This driven element may either be a disk for rotary movement or a cylinder for translatory movements.

Table 9.6 Technical properties of the Elliptec-motor X15G [4].

Symbol	Parameter	Values			Unit
		min.	typ.	max.	
v_0	idle mode	300	350	550	mm/s
a	acceleration ftom 0 up to max. velocity at $m = 1,5g$		5		ms
F_{H}	holding force in zero-current case	0,5	0,8	1,2	N
F_{B}	blocking force	200	300	500	mN
F_{100}	actuating force at 100 mm/s	100	200	350	mN
F_{200}	actuating force at 200 mm/s	25	100	200	mN
M_{B}	blocking-toraue at 20mm ⌀ of teh driven rotor	2	3	5	mNm
M_{100}	torque at 20mm ⌀ and 95,5 rpm	1	2	3,5	mNm
M_{200}	torque at 20mm ⌀ and 191 rpm	0,25	1	2	mNm
f_{fw}	frequency forward mode at v_{max}	77	81	84	kHz
f_{bw}	frequency backward mode at v_{max}	93	98	108	kHz
f_{r}	frequency resolution		0,2	0,6	kHz
f_{D}	temperature-depending frequency drift	35	50	70	Hz/°C
r_{s}	positioning accuracytre	var. 5 bis 100			μm
P_{p0}	energy of driver electronics		1,8		W
I_{driver}	max. current of driver electronics at 5 V	300	450	600	mA
U_{motor}	motor voltage (peak-to-peak)	5	7	10	V

For control of the motor just a single frequency f_1 has to be provided. Due to the actuators geometrical design the elliptic movement changes its direction at a second frequency f_2 resulting in a movement change on the drive side.

A typical length of the actuator is 20 mm. The resonator shows a height and width of 3 mm. The weight is 1,2 g. The major technical specifications of the Elliptec-drive are given in table 9.6. Due to its components and the very compact driver electronics it can be obtained for little costs at low quantities.

Ultrasonic Actuators with Circular Resonators

As mentioned before beside actuators providing standing waves another group of actuators based on traveling waves exists. The traveling wave actuators known best are circular in their design. The first actuator according to this principle has been built in 1973 by SASHIDA [261]. Traveling waves actuators count to the group of ultrasonic actuators as their control frequencies typically lie in between 20 kHz up to 100 kHz liegen.

This section is reduced to the presentation of ring-shaped traveling wave actuators with a bending wave. Other design variants for linear traveling wave actuators can be found in the corresponding literature [54, 92, 93].

Figure 9.43 shows an actuator's stator made of piezoelectric elements. They have an alternating polarization all around the ring. The stator itself carries notches actually enabling the formation of the rotating traveling wave.

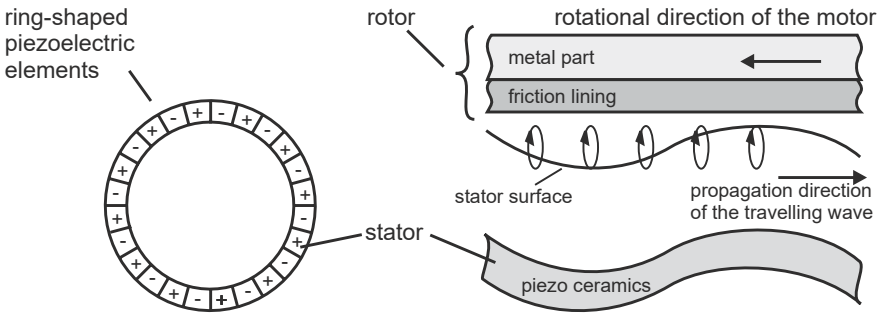


Fig. 9.43 Piezoelectric traveling wave motor: left: stator disk with piezoelectric elements. right: schematic view of the functionality of a ring-shaped piezoelectric traveling wave motor [116]
.

Each point on the surface of the stator performs a local elliptic movement (trajectory). This movement is sketched schematically in figure 9.43 These individual elliptic movements overlay to a continuous wave on the stator. With a frictional coupling this movement is transferred on the rotor, resulting in a rotation. The contact between stator exists with the same number of contact points anytime during operation.

The movement equation of the traveling wave actuator is given by

$$u(x,t) = A\cos(kx - \omega t) \tag{9.68}$$

By reshaping it the following form results in

$$u(x,t) = A(\cos(kx))(\cos(wt)) + A(cos(kx - \pi/2))(cos(kx + \pi/2)) \tag{9.69}$$

The second term of equation 9.69 includes important information for the control of traveling wave actuators. A traveling wave can be generated by two standing waves being spatially and timely different. Within typical realization the spatially difference of $x_0 = \lambda/4$ unis chosen, and a time phase lag of $\Phi_0 = \pi/2$. The usage of two standing waves is the only practical possibility for generating a traveling wave. The direction of rotor-movement can be switched by changing the phase lag from $+\pi/2$ to $-\pi/2$ [81, 95, 99, 261].

Figure 9.44 shows the practical realization of a traveling wave motor.

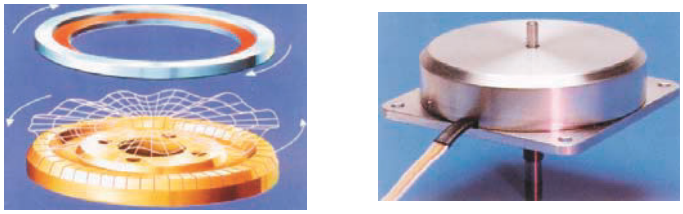


Fig. 9.44 Realization of a traveling wave motor by Daimler-Benz AG [279].
.

Big advantages of a traveling wave motor are the high torques possible to achieve at low rotational speeds. It has a low overall size and is of little weight. This enables a very thin design as shown in figure 9.44 . In passive mode the traveling mode motor has a high locking torque of several Nm.

Other advantages are given by the good control capabilities, the high dynamics, and robustness against electromagnetic noise such as the silent movement [282]. Typical applications of traveling wave actuators are autofocus-functions in cameras.

Uchinomotor

The piezoelectric actuator of K. UCHINO [260, 261] has a geometrically very compact design (fig. 9.45). It is made of a metallic tube which was flattened on two sides to provide two orthogonal surfaces. Two rectangular piezoelectric elements are attached to these surfaces.

By actuating plate X the metallic tube experiences a bending torque in the axial direction of x. Due to the asymmetrical mass distribution resulting from the plate Y on the cylinder, a second bending movement in y direction is induced. This move-

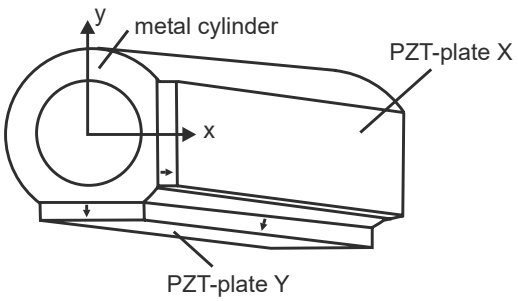


Fig. 9.45 Schematic view of the piezoelectric ultrasonic actuator by UCHINO [260, 261].

ment is of identical phase with the movement in x-direction. The resulting combined movement is elliptic. If another rod is placed in the metallic tube, it is driven by the combined movement in axial translational and rotational direction. To change the direction of movement the plate Y is actuated.

The actuator is extremely small. Its typical diameter is given with 2.4 mm at a length of 12 mm. It excels by its simple design and easy and cheap manufacture. The controlling frequency lies at 61 kHz. In idle mode the achieved rotational speed is 1800 rpm and a torque of 1.8 mNm. The translational velocity can be varied between $1\,\mu m/s$ to $10\,mm/s$ at a resolution of $0{,}5\,\mu m$. The controlling voltages are 80 V.

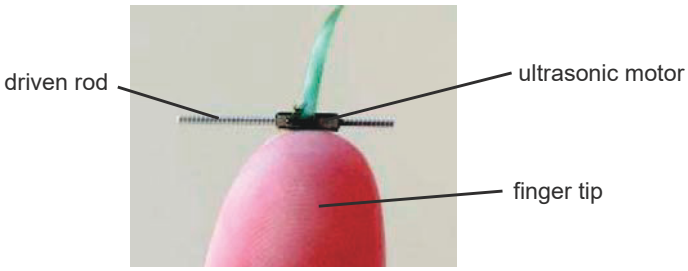


Fig. 9.46 Photo of the implementation of the piezoelectric UCHINO-motor by Squiggle Motors [238].

A very small variant of this actuator has the geometrical dimensions of 1,5 x 1,5 mm^2 at a length of 6 mm and is manufactured by Squiggle Motors (fig. 9.46). It is the smallest linear actuator available commercially at the time of print.

Piezoelectric Stepper Motors

Another interesting design can be found with the actuator PI Nexline. It combines the longitudinal effect with the piezoelectric shear effect, resulting in a piezoelectric stepper motor.

The principle design is sketched in figure 9.47.

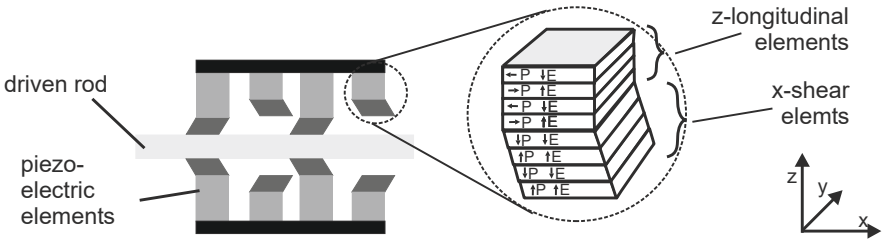


Fig. 9.47 Piezoelectric stepper motor using the shear effect and the longitudinal effect [134].

The movement of the motor is similar to the inchworm-principle. Drive- and release-phases of the piezoelectric elements produce a linear movement of the driven rod. The piezoelectric longitudinal elements generate the clamping force in z-direction, the shear elements rotated by 90° a translational movement in y-direction is possible too.

The advantage of this design is given by the high positioning resolution. Over the whole displacement of 20 mm a resolution of 0.5 nm can be achieved. The stepping frequency is given by - dependent on the control - up to 100 Hz and enables, depending on its maximum step-width, velocities of up to 1 mm/s. The step-width can be chosen stepless between 5 nm and 8 μm. The intended position can be achieved either closed-loop or open-loop controlled. For the closed-loop control a linear encoder has to be added to the motor. In open-loop control the resolution can be increased to 0.03 nm in a high resolution dithering mode.

The actuator can generate push- and pull-forces of 400 N maximum. The self-locking reaches up to 600 N. The typical driving voltage is 250 V. The specifications given above are based on the actuator N-215.00 Nexline® of the company Physik Instrumente (PI) GmbH & Co. KG (www.pi.ws) [134]. Beside the impressive forces and positioning resolutions which can be achieved, these actuators have a high durability compared to other designs of piezoelectric actuators, as no friction happens between moving parts and stator.

An application of the stepper motor is given in figure 9.48. Six of above drives are used within in a hexapod parallel kinematic design, which can be used to position high loads in the presence of strong magnetic fields.



Fig. 9.48 Application of the piezoelectric stepper motor in a hexapod [134].

9.4.3 Design of Piezoelectric Actuators for Haptic Systems

Within the preceding section the basic designs of piezoelectric actuators have been discussed and special variants were shown. This section transfers this knowledge about the design of piezoelectric actuators focussing on haptic applications now.

First of all the principle approach for designing is sketched. Hints are given about which designs are appropriate for which applications. Afterward three tools for practical engineering are shown: description via electromechanic networks, analytical formulations and finite-element simulations

9.4.4 Procedure for the Design of Piezoelectric Actuators

Figure 9.49 gives the general procedure for the design of piezoelectric actuators.

The choice of a general design based on those shown in the prior section is largely dependent on the intended application. For further orientation figure 9.50 shows a decision tree for classifying the own application.

The following paragraph describes the appropriate designs for specific application classes according to this scheme. The list has to be regarded as a point for orientation, but it does not claim to be complete. The creativity of an engineer will be able to find and realize other and innovative solutions beside those mentioned here. Nevertheless especially for the design of tactile devices some basic advice can be given:

1 Braille displays have to act against the finger's force. At typical resolutions this requires forces in the area of mN at displacement of around 100 μm. The requirements on the dynamics are in the lower range of several Hertz. The smallest resolution of a pixel has to be in the area of 1 x 1 mm^2 and is defined by the

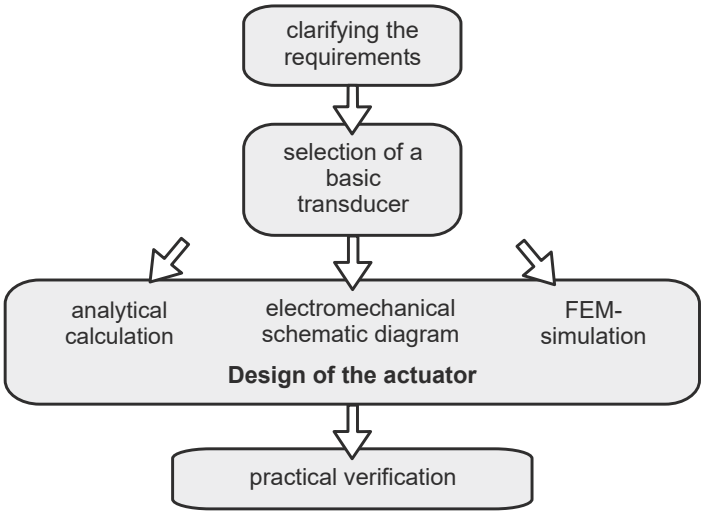


Fig. 9.49 Procedure of designing piezoelectric actuators.

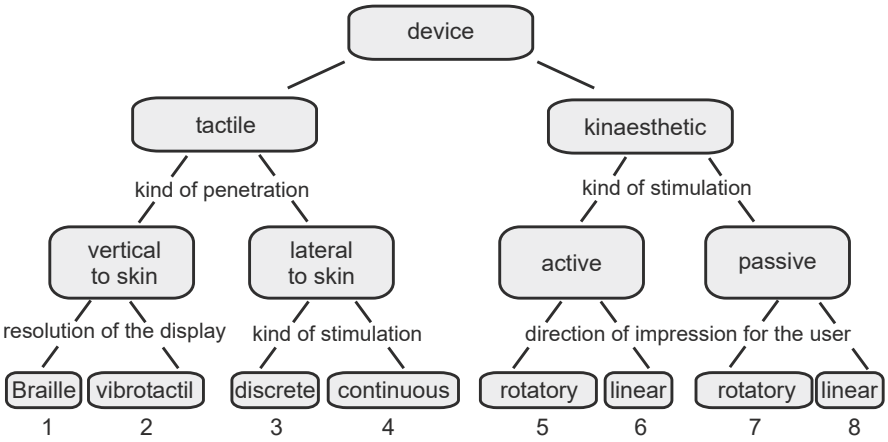


Fig. 9.50 Decision tree for the selection of a type of piezoelectric actuator.

lowest resolution at the finger's tip. Looking at the force-amplitude diagram of figure 9.51 bending-actuators fit well to these requirements.

2 In comparison to Braille-displays there are vibrotactile displays, which need higher frequencies and smaller displacements and forces to present a static shape to the user. With the diagram of figure 9.51 especially bending disc or staple actuators would be appropriate to fulfil these requirements, although there are overpowered concerning the achievable forces.

3,4 Such displays are subject to current research and are not yet applied in broad range. Their design is typically based on bending actuators, as shear the skin requires less force than forces acting normal to the skin by generating a comparable perception.

5-8 In contrast to tactile displays the actuator selection for kinaesthetic systems is more dependent on the actual application. Forces, displacement and degrees of freedom influenced by the kinematics alter in wide ranges. Additionally the actuator's volume may be a criterion for selection. Figure 9.51 gives an overview about piezoelectric actuation principles and has to be interpreted according to the specific kinaesthetic problem at hand. Generally speaking ultrasonic piezoelectric actuators are usually the matter of choice for kinaesthetic devices, although they have to be combined with a closed-loop admittance control.

Figure 9.51 gives an overview about piezoelectric actuation principles and has to be interpreted according to the specific kinaesthetic problem at hand. Generally speaking ultrasonic piezoelectric actuators are usually the matter of choice for kinaesthetic devices, although they have to be combined with a closed-loop admittance control.

Additional reference for actuator selection can be found in chapter 9.4.2 are suitable for haptic applications, but still need some care in their usage due to high voltages applied and their sensitivity on mechanical damage. This effort is often rewarded by piezoelectric actuation principles, which can be combined to completely new actuators. And the only thing required is some creativity of the engineer.

After choosing the general actuator the design process follows next. For this purpose three different methods are available, which are presented in the following and discussed in their pros and cons. In addition some hints on further references are given.

9.4.4.1 Methods and Tools for the Design Process

There are three different engineering tools for the design of piezoelectric actuators:

- Description via the aid of electromechanical concentrated networks
- Analytical descriptions
- Finite elements simulations

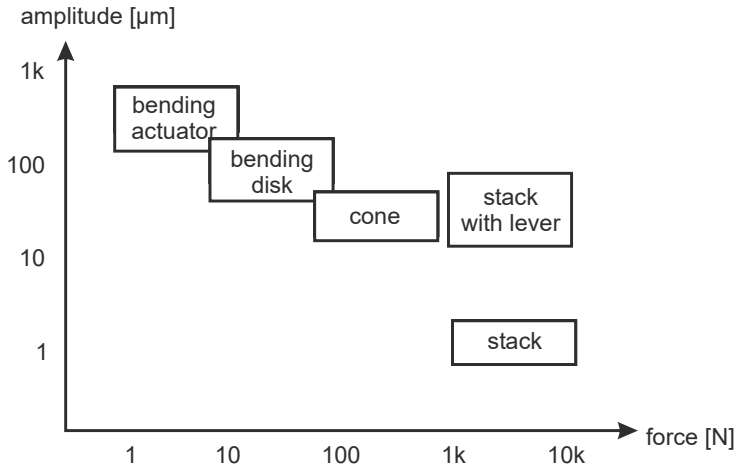


Fig. 9.51 Force-amplitude-diagram for the classification of the piezoelectric actuating types.

Description via the Aid of Electromechanical Concentrated Networks

The piezoelectric basic equations from paragraph 9.4.1.1 on page 244 are the basis for the formulation of the electromechanic equivalent circuit of a piezoelectric converter.

The piezoelectric actuator can be visualized as an electromechanical circuit. Figure 9.52 shows the converter with a gyratory coupling (a), alternatively a transformatory coupling (b) is possible too (see appendix 17).

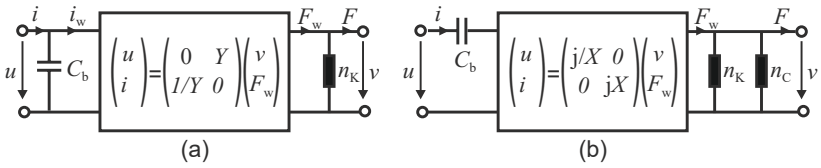


Fig. 9.52 Piezoelectric actuator as a electromechanical schematic diagram (a) gyratory and (b) transformatory combination [158].

For the gyratory coupling the equations 9.70 to 9.73 summarize the correlations for the calculation of the values for the concentrated elements. They are derived from the constants e, c, ε ssuch as the actuator's dimensions l and A [158].

$$C_b = \varepsilon \cdot \frac{A}{l} = (\varepsilon - d^2 \cdot c)\frac{A}{l} \qquad \text{with} \quad v = 0 \tag{9.70}$$

$$n_K = \frac{1}{C} \cdot \frac{l}{A} = s \cdot \frac{A}{l} \qquad \text{with} \quad U = 0 \tag{9.71}$$

$$Y = \frac{1}{e} \cdot \frac{l}{A} = \frac{s}{d} \cdot \frac{l}{A} \tag{9.72}$$

$$k^2 = \frac{e^2}{\varepsilon \cdot c} = \frac{d^2}{\varepsilon \cdot c} \tag{9.73}$$

With the piezoelectric force constants

$$e = d \cdot c = \frac{d}{s} \tag{9.74}$$

Which makes for the transformatory coupling

$$X = \frac{1}{\omega C_b \cdot Y} \qquad \text{und} \qquad n_C = Y^2 \cdot C_b \tag{9.75}$$

Figure 9.53 shows the sketch of a element Δx taken out of a piezoelectric bimorph bending actuator (dimensions Δl x Δh x Δb) as a electromechanical equivalent circuit.

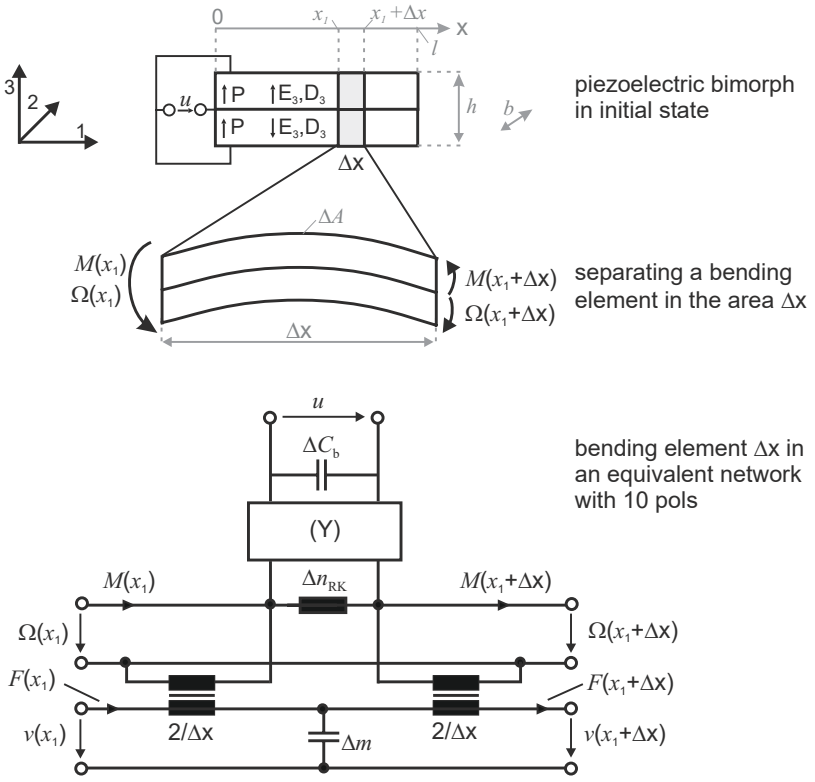


Fig. 9.53 Piezoelectric bimorph bending element in a electromechanical schematic view in quasistatic state [158].

It is:

$$C_b = 4\varepsilon_{33}^T(1-k_L^2)\frac{b\cdot\Delta x}{h}$$

$$\Delta n_{RK} \approx 12 s_{11}^E \frac{(\Delta x)^3}{b\cdot h^3}$$

$$\frac{1}{Y} = \frac{1}{2}\frac{d_{31}}{s_{11}^E}\frac{b\cdot h}{\Delta x}$$

The piezoelectric lossless converter couples the electrical with the mechanical rotatory coordinates first, which are torque M and angular velocity Ω. To calculate the Force F and the velocity v an additional transformatory coupling between rotatory and translatory mechanical network has to be introduced. As a result the complete description of a sub-element Δx of a bimorph is given in a ten-pole equivalent circuit.

Analytical Calculations

A first approach for the design of piezoelectric actuators is given by the application of analytical equations. The advantage of analytical equations lies in the descriptive visualization of physical interdependencies. The influence of different parameters on a target value can be derived directly from the equations. This enables a high flexibility in the variation of e.g. dimensions and material properties. Additionally the processing power for the solution of equations can - compared to simulations - be neglected.

A disadvantage of analytical solution results from the fact, that they can only be applied to simple and frequently only symmetrical geometrical designs. Although already limited to such designs even simple geometries may result in very complex mathematical descriptions requiring a large theoretical background for their solution.

The following collection gives relevant literature for the familiarization with specific analytical challenges being faced with during the design of piezoelectric actuators:

- Very compelling and complete works on the design of piezoelectric actuators are [260, 261, 263, 262].
- The theory of the piezoelectric effect and piezoelectric elements are discussed in [29, 104, 105].
- The mathematical description of traveling wave actuators can be found in [214, 294].
- The contact behavior between stator and rotor with piezoelectric multilayer bending-actuators is analyzed in [80, 248, 278, 297]
- In [7, 8] the static and dynamic behaviour of multilayer beam bending actuators is described.
- The description of the mechanical oscillations for resonance shapes is discussed in [16, 48, 55, 140, 228, 280] elaborately.

Finite Element Simulation

The application of both approaches given before is limited to some limited geometrical designs. In reality complex geometrical structures are much more frequent which can not be solved with analytical solutions or mechanical networks. Such structures can be analyzed according to the method of finite element simulation (FEM).

For the design of piezoelectric actuators the use of coupled systems is relevant. One example of a FEM-simulation for a piezoelectric travelling wave motor is shown in figure 9.54

For the solution of the differential equations the market offers a reasonable number of simulation tools. At this position only some shall be named explicitly

- ANSYS (www.ansys.com)
- ATILA (www.cedrat.com)

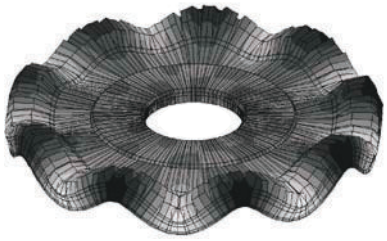


Fig. 9.54 FEM-simulation of the oscillation shape of the stator of a piezoelectric travelling wave motor (view highly exaggerated) [282].
.

- Comsol Multiphysics (www.femlab.de)
- ProMechanica (www.ptc.com) oder
- GiD (http://gid.cimne.upc.es).

The simulation tools are subject to costs, but free test-versions limited in the number of nodes or the availability certain functions shapes can also be found. Literature for the start is offered in form of books like “FEM für Praktiker” [179, 241] such as manifold documentation and handbooks of the individual software.

9.4.5 Piezoelectric Actuators in Haptic Systems

Piezoelectric actuators are among the most frequently used actuation principles in haptic systems. The designs shown before can be optimized and adapted for a reasonable number of applications. One of the most important reasons for their usage is their effectiveness at a very small required space, which is identical to a high power density. To classifiy the realized haptic systems a division into tactile and kinaesthetic is done for the following paragraphs.

9.4.5.1 Piezoelectric Actuators for Tactile Systems

For the design of any tactile systems the application area is of major importance. The bandwidth ranges from macroscopic table top devices, which may be used for embossed printings in Braille being placed below a PC-keyboard, up to highly integrated systems, which may be used for mobile applications. Especially for latter use, the requirements on volume, reliable and silent operation, but also on low weight and energy consumption are enormous. The following examples are structured into two subgroups. Each of them addresses one of two directions of the penetration of the skin: lateral and normal.

Tactile Displays with Normal Penetration

Braille Devices

A Braille character is encoded by a dot pattern formed by embossed points on a flat surface. By touching this pattern made of eight dots (two columns with four rows of dots each) combinations of up to 256 characters can be addressed. Since the seventies reading tablets for visually handicapped people are developed which are capable to present these characters with a 2x4 matrix of pins. The most important technical requirements are a maximum stroke of $0.1 - 1\,\mathrm{mm}$ and a counter force of $200\,\mathrm{mN}$. Already early in this development electromagnetic drives have been replaced by piezoelectric bimorph bending actuators. These actuators enable a thinner design, are more silent during operation and faster. At typical operating voltages of $\pm 100 - 200\,\mathrm{V}$ and a nominal current of $300\,\mathrm{mA}$ they additionally need less energy than the electromagnetic actuators used before. Figure 9.55 shows the typical design of a Braille character driven by a piezoelectric bimorph actuator. A disadvantage of this system is the high price, as for 40 characters with eight elements each all together 320 bending actuators are needed. Additionally they require still a large volume as the bending elements have to show a length of several centimeters to provide the required displacements. This group of tactile devices is part of the shape-building devices. The statically deflected pins enable the user to detect the displayed symbol.

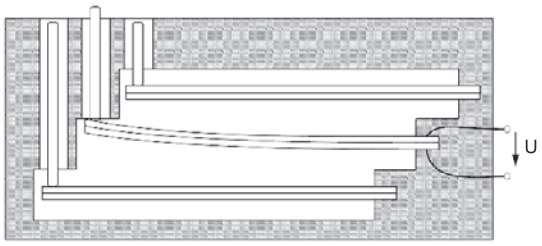


Fig. 9.55 Schematic setup of a Braille row with piezoelectric bending actuators

Vibrotactile Devices

With vibrotactile devices the user does not detect the displacement of the skin's surface but the skin itself is put into oscillations. At smaller amplitude the sensation is similar to the static elongation. The general design of vibrotactile displays equals an extension of the Braille-character to an N x N matrix which is actuated dynamically. The tactile image generated is not perceived by the penetration depth but by the amplitude of the oscillation [107]. Another impact factor is the oscillation frequency, as the tactile perception depends extremely on the frequency. With the knowledge of these interdependencies optimized tactile displays can be built generating a very well perceivable stimulation of the receptors. Important for displays according to this approach is a large surface, as movements performed by the own finger disturb the perception of the patterns.

The "Texture Explorer" is designed as a vibrating 2x5 pin array [106]. It is used for the research on the perception of tactile stimulation, such as the overlay of tactile stimulations with force-feedback. The surfaces displayed change according to their geometry and their roughness within the technical limits of the device. Figure 9.56 shows a schematic sketch of the actuator array. The contact-pins have a size of $0.5 \times 0.5\ \mathrm{mm}^2$ with a point-to-point distance of 3 mm. Each pin is actuated separately by a bimorph bending actuator at a driving voltage of 100 V and a frequency of 250 Hz. The maximum displacement of these pins with respect to surface level is $22\,\mu$m and can be resolved down to $1\,\mu$m resolution.

An even more elaborate system is based on 100 individually controlled pins [244]. It can be actuated dynamically in a frequency range of $20 - 400$ Hz. Figure 9.57 shows a schematic sketch. 20 piezoelectric bimorph bending actuators (PZT-5H, Morgan Matrinic, Inc.) in five different layers one above the other are located in a circuit around the stimulation area. Each bending actuator carries one stimulation pin, which is placed 1 mm above the surface in idle state. The pins have a diameter of 0.6 mm and are located equally spaced at 1 mm distance. At a maximum voltage of ± 85 V a displacement of $\pm 50\,\mu$m is achieved. A circuit of equally high passive pins is located around the touch area to mark the borders of the active display.

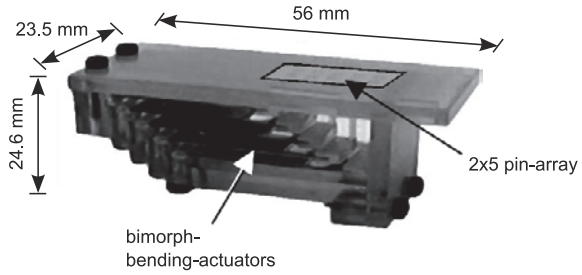


Fig. 9.56 Schematic setup of the "Texture Explorer" [106]

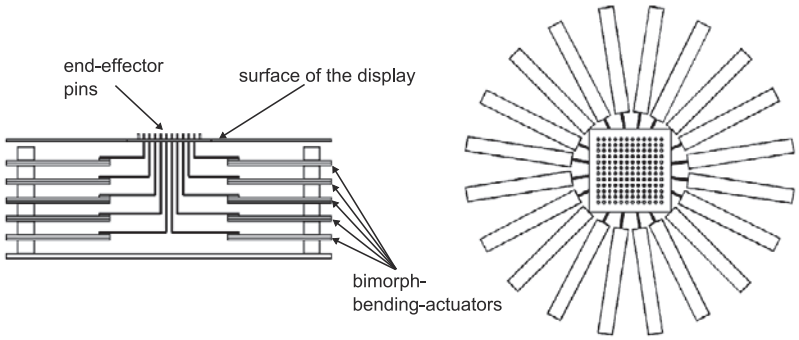


Fig. 9.57 Schematic setup of the 100-pin-array [244]

Another even more compelling system can be found in [149] (fig. 9.58). The very compact 5 x 6 array is able to provide static and dynamic frequencies of up to $\approx 500\,\text{Hz}$. Once again, piezoelectric bending actuators are used to achieve a displacement of $700\,\mu\text{m}$. However the locking-force is quite low with a maximum of 60 mN.

Ubi-Pen

The "Ubi-Pen" is one of the highest integrated tactile systems. Inside of a pen both components, a spinning disc motor and a tactile display, are assembled. The design of the tactile display is based on the "TULA35" ultrasonic linear drive (Piezoelectric Technology Co). The schematic sketch of the design is given in figure 9.59. The actuator is made of a driving component, a rod and the moving part. The two piezoelectric ceramic discs are set to oscillation resulting into the rod oscillating up- and downwards. The resulting movement is elliptical. To move upwards, the following procedure is applied: in the faster downward movement the inertial force is excelled by the speed of the operation and the element remains in the upper position. Whereas

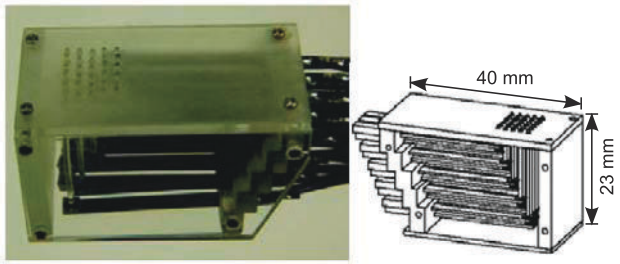


Fig. 9.58 Schematic setup of the 5x6-pin-array [149]

in the upwards movement the speed is controlled slow enough to carry the moving part up by the frictional coupling between moving element and central rod. The actuator discs have a diameter of 4 mm and a height of 0.5 mm. The rod has a length of 15 mm and a diameter of 1 mm. It can be used as contact-pin directly. The actuator's blocking force is larger than 200 mN and at a control frequency of 45 kHz velocities of 20 mm/s can be reached.

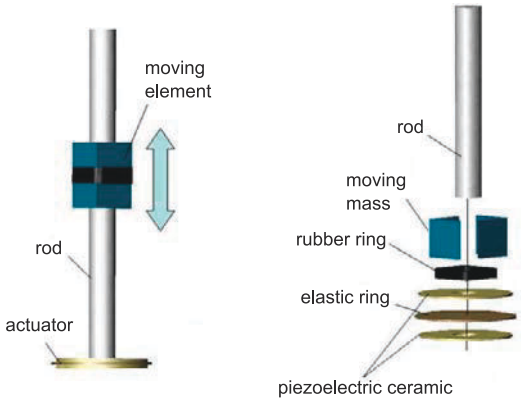


Fig. 9.59 Schematic setup of the ultrasonic motor "TULA35" [152]

Figure 9.60 shows the design of a 3 x 3 pin array. Especially the very small size of the design is remarkable: all outer dimensions have a length of 12 mm. The pins are distributed in a matrix of 3 mm. On the area of 1.44 cm^2 nine separate drives are located. To achieve such a high actuator density the lengths of the rods have to be different, allowing the moving parts to be placed directly next to each other. If this unit is placed at the upper border of the display all pins move in- respectively out-of

the plane. The weight of the whole unit is 2.5 g. When the maximum displacement of 1 mm is used, a bandwidth of up to 20 Hz can be achieved.

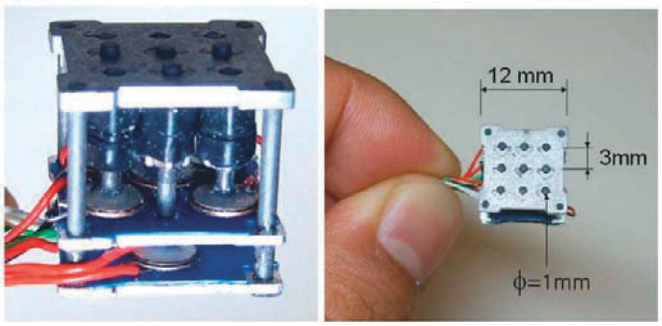


Fig. 9.60 Tactile 3x3-pin-array [152]

The integration in a pen including another vibratory motor at its tip is shown in figure 9.61. This additional drive is used to simulate the contact of the pen with a surface. The whole pen weights 15 g.

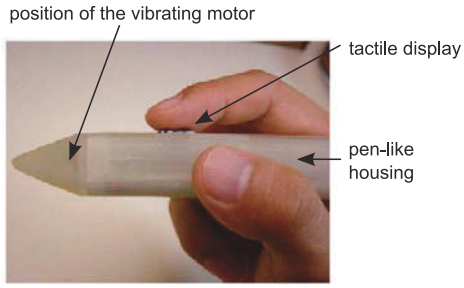


Fig. 9.61 Prototype of the "Ubi-Pen" [151]

The Ubi-Pen provides surface structures such as roughnesses and barriers or other extreme bumpy surfaces. To realize this, vibrations of the pins are superimposed with the vibratory motor. If the pen is in contact with a touch-sensitive surface (touch-panel), the shown graphical image may be displayed in its gray scale values by the pins of the tactile display. The system has been used for a number of tests for recognition of information displayed in different tactile modalities [150]. The results are very good with a mean recognition rate of 80 % with untrained users.

Tactile Displays with Lateral Penetration

Discrete Stimulation

The concept of discrete stimulation is based on an excitation lateral to the skin's surface ("laterotactile display") [91]. On the left hand in figure 9.62 the schematic sketch of a one-dimensional array of actuators is given. An activation of a piezoelectric element results in its elongation and a deformation of the passive contact comb (crown). If the skin of a touching finger is stretched by this movement a contact point is perceived. Figure 9.62 shows a 2-dimensional display on the right. With the extension from 1D to 2D array it is necessary to consider the more complex movement patterns resulting from it. A more deep analysis of the capabilities of such a system is found in [161], proving the applicability of a laterotactile display as a virtual 6-point Braille device.

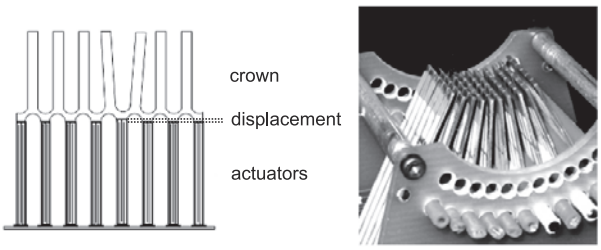


Fig. 9.62 1D array of the "laterotactile display" [91] and 2D STReSS^2 display [161]

Continuous Stimulation

The transfer from the discrete points to a piezoelectric traveling wave drive is shown in [22]. The touching finger faces a closed and continuous surface. Due to this design the tactile display itself becomes less sensitive in its performance to movements of the finger. With the contact surface beyond the skin being excited as a standing wave, the user perceives a surface texture. With relative movement between wave and finger even a roughness can be perceived. By modifying the shape of the travelling wave a simulation of a touch-force perceivable by the moving finger can be achieved. Figure 9.63 shows the schematic sketch of the contact between finger and traveling wave, such as the corresponding movement direction. In a test bed application [22] the stator of a travelling actuator USR60 from Shinsei has been used. This actuator provides a typical exploration speed by its tangential wave speed of 15 cm/s and forces up to ≈ 2 N. This system enables to generate continuous and braking impressions by the change of wave shapes. An additional modulation of the ultrasonic signals with a low frequency periodical signal generates the sensation of

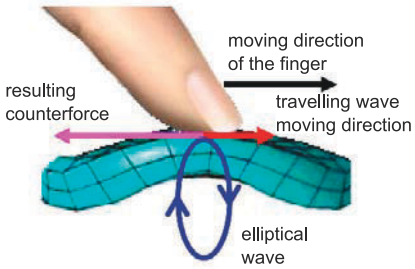


Fig. 9.63 Contact area of finger and travelling wave [22]

surface-roughnesses. Actual research is performed on the design of linear ultrasonic travelling wave displays.

9.4.5.2 Piezoelectric Actuators for Kinaesthetic Systems

Piezoelectric actuators used in kinaesthetic systems are usually part of active systems (in a control-engineering sense). The user interacts with forces and torques generated by the actuator. A classic example is a rotational knob which is actuated by a travelling wave motor. With passive systems the actuator is used as a switching element, which is able to consume power from an actuator or user in either time-discrete or continuous operation. Examples are breaks and clutches.

Active Kinaesthetic Systems

Piezoelectric traveling wave actuators show a high torque to mass ratio if they are compared to other electrical actuation principles. They are predestined for the use in applications with a high torque at small rotational velocity, as they do not need an additional bearing or other transmission ratios. Kinaesthetic systems require exactly these properties. A very simple design can be used to build a haptic knob: a rotationally mounted plate with a handle attached for the user is pressed on the stator of a traveling wave piezoelectric motor. A schematic sketch of the critical part is shown in figure 9.64. Due to the properties specific to the traveling wave motor the rotational speed of the rotor can be adjusted easily by increasing the wave's amplitude w at the stator. As this actuation principle is based on a mechanical resonance mode, it is actuated and controlled with frequencies nearby its resonance. Coincidentally, this is the most challenging part of the design, as the piezoelectric components show a very nonlinear behavior at the mechanical stator-rotor interface. Hence, the pro-

cedures for its control and its electronics have a large influence on its performance.

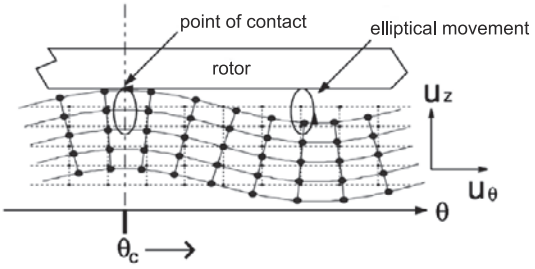


Fig. 9.64 Characteristics of the stator's traveling wave beneath the rotor [72]

Figure 9.65 shows the speed vs. torque characteristics for different wave amplitudes *w* of the actuator [43]. The torque is highly dependent on the actual rotational speed and the amplitude. By monitoring the speed and controlling the phase and the wave amplitude, the system can be closed-loop controlled to a linear torque-displacement characteristic. In this mode a maximum torque of ≈ 120 mNm can be achieved with this example. A deeper discussion of the phase control for a piezoelectric travelling wave motor according to this example is given in [72].

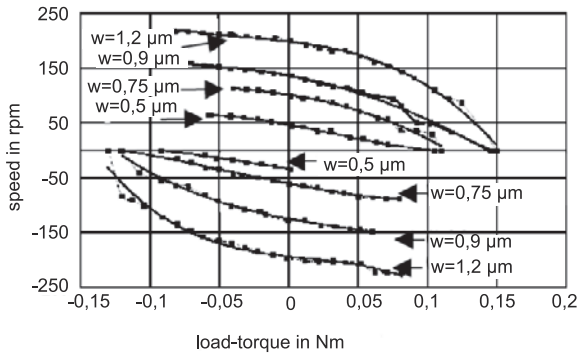


Fig. 9.65 Speed / load torque-characteristics for different wave amplitudes [43]

A specialized design of such a device is used in neurological research for the application with magneto-resonance tomography [59]. To close-loop control the admittance of the device the torque has to be measured. In the specific application near a large magnetic field glass fibers are used, measuring the reflection at a bending

polymer body. Preventing disturbance from the device on the switching gradients of MRI and vice versa the device has been designed from specific non-conductive materials. It was based on a traveling wave motor in a special MR-compatible version of the "URS60" from Shinsei.

Hybrid Systems

Another class of kinaesthetic systems are so called hybrid systems. If there is the need of generating a wide bandwidth of forces and elongations with a single device, there is no actuator fulfilling all requirements alone. Due to this reason several hybrid systems are designed with two (or more) components complementing each other. A typical example is the combination of a dynamic drive with a brake, latter is used to provide large blocking torques. As seen in the above paragraph the closed-loop control of a traveling wave motor's impedance is a challenging task. A simple approach to avoid this problem is the combination of a traveling wave actuator with a clutch. The main difference between a traveling wave actuator and other types of actuators is given by its property to rotate with a load-independent velocity. Providing a certain torque this system is accompanied with a clutch. Other designs add a differential gear or a break. Such a system is presented in [37]. If the system experiences a mechanical load, the operation of the break is sufficient to increase friction or even block the system completely. Consequently, the system provides the whole dynamic range of the traveling motor in active operation, whereas passive operation is improved significantly by the break. Due to the simple mechanical design and the reduction of energy consumption such systems are suitable for mobile applications, as well.

Passive Kinaesthetic Systems

Objects can be levitated by standing acoustic waves. To achieve this, an ultrasonic source has to provide a standing wave first. Within the pressure nodes of the wave a potential builds up, attracting a body placed nearby the node. The size of the object is important for the effect to take place, as with a large size the influence on the next node may be too high. A system based on this principle is described in [145]. It shows a design of an exoskeleton in form of a glove with external mechanical guides and joints. The joints are made of piezoelectric clutches with a schematic design shown in figure 9.66. In their original state both discs are pressed together by an external spring generating a holding torque. With the vibrator being actuated the levitation mode is achieved between rotor and stator creating a gap h. This reduces the friction drastically allowing to turn both disks almost free from each other.

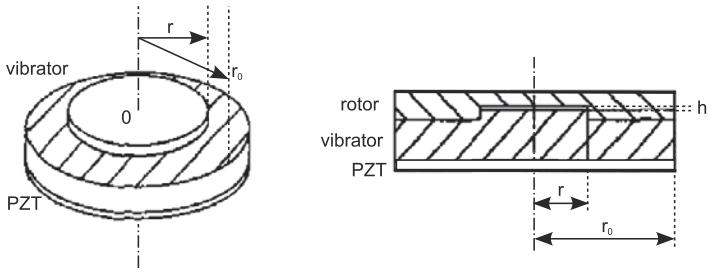


Fig. 9.66 Schematic setup of the "levitation-clutch" [145].

9.4.5.3 Summary

Tactile systems are distinguished according to their direction of movement. With a normal movement into the direction of the skin additional differences are made between passive systems simulating a more or less static surface by their pins, and active systems - so called vibrotactile systems - providing information by a dynamic excitement of the skin's surface. The user integrates this information into a static perception. The advantages of this approach are given by the reduced requirements on the necessary force and displacements, as the dynamical perception of oscillations is higher than the perception of static slow-changing movements. When the display is not fixed to the finger however its fast movements will be a problem. With a static display in a fixed frame the user is able to repeatedly touch the display, increasing the dynamics of the haptic impression by own movements. With a dynamic display this interaction does not work as well anymore as periods of oscillations from the vibrating elements are lost.

Another alternative are tactile systems with a lateral movement of the skin. With an appropriate control the human can be "fooled" to feel punctual deformations analog to an impression of a normal penetration. Systems with a closed surface are very comfortable variants of such displays, but their dynamic control is demanding for finger positions moving across larger surfaces. Typically today's solutions show smaller contact areas than with other variants, as the actuator elements can not be placed as close together as necessary.

Kinaesthetic (force-feedback) systems can be distinguished in active and passive systems according to a control-engineering sense. Active systems are able to generate counter forces and supportive forces. The spectrum of movements is only limited by the degrees-of-freedom achieved by the machanical design. A stable control for active system tends to become very elaborate due to required measurement technology and complex control algorithm of sufficient speed. As with all active systems a danger remains: an error of its functionality of the system may harm the user. This danger increases for piezoelectric actuators, as the available forces and torques

are high. Passive systems with breaks and clutches enable the user to feel the resistance against their own movement as reactive forces. These designs are simpler to build and less dangerous by definition of passivity. General disadvantages of passive systems can be found in their high reaction times, the change of their mechanical properties in long-time applications and their comparably large volume. Hybrid systems combining both variants - usually including another actuation principle - may enlarge the application area of piezoelectric actuators. Although the mechanical design increases in volume and size, the requirements on control may become less and large holding forces and torques can be achieved with low power consumption. From a standpoint of haptic quality they are one of the best actuator solutions for rotating knobs with variable torque/angle characteristics available today.

9.5 Electrostatic Actuators

MARC MATYSEK

Electrostatic transformers are part of the group of electric transformers, such as the piezoelectric actuators, too. Electric transformers show a direct coupling between the electrical value and the mechanical value. This is contrary to electrodynamic and electromagnetic actuators, which show an intermediate transformation into magnetic values as part of the actuation process. In principle the transformation may be used in both directions. Hence, all these actuators can be used as sensors, as well. Electrostatic field actuators are utilized due to their simple design and low power consumption. As a result from the technical progress of micro-engineering, the advantages of integrated designs are fully utilized. Especially for miniaturized systems the electrostatic field actuators gain increased importance compared to all other actuator principles. This is even more surprising as their energy density is significantly lower in macroscopic designs. But during miniaturization the low efficiency factor and the resulting power loss and heat produced become limiting factors for magnetic actuators [159]. An important subgroup of electrostatic field actuators is given by a solid state actuators, with an elastomeric dielectric. It has a high breakdown field strength compared to air, builds the substrate of the electrodes and can simultaneously provide an isolating housing, too. Beside the classic field actuators mentioned above, electro rheological fluids are part of the electrostatic actuators, as well. With these actuators an electric field of an arbitrary external source results in a change of the physical (rheological) properties of the fluid.

9.5.1 Definition of the Electric Field

The following paragraphs define the electric field and relevant variables for the design of electrostatic actuators.

9.5.1.1 Force on Charge

The magnitude of a force F acting on two charges Q_1 and Q_2 in a distance r is given by COULOMB'S-law (equation (9.76)).

$$F = \frac{1}{4\pi\varepsilon_0}\frac{Q_1 Q_2}{r^2} \tag{9.76}$$

9.5.1.2 Electric Field

The electric field E describes the space where these forces are present. The field strength is defined as the relation of the force **F** acting on the charge in the field and the charge's magnitude Q.

$$\mathbf{E} = \frac{\mathbf{F}}{Q} \tag{9.77}$$

The charges cause the electric field; the forces on the charges within an electric field are the effect. Cause and effect are proportional. With the electric constant $\varepsilon_0 = 8,854 \cdot 10^{-12}$ C/Vm within vacuum and air equation (9.78) results:

$$\mathbf{D} = \varepsilon_0 \mathbf{E} \tag{9.78}$$

The electric displacement field **D** describes the ratio of the bound charges and the area of the charges. The direction is given by the electric field pointing from positive to negative charges. If the electric field is filled with an insulating material (dielectric), the electric displacement field is bound partly due to the polarizing of the dielectric. Accordingly, the field strength drops from E_0 to E (with still the same electric displacement field). Consequently, the ratio of the weakened field depends on the polarizability of the dielectric and is called "permittivity" $\varepsilon_r = E_0/E$.

9.5.1.3 Capacity

The electrical capacity is defined as the ratio of charge Q on each conductor to the voltage U between them. A capacitor with two parallel-plates charged contrary with a surface of the plates A and a fixed distance d shows a capacity C depending on the dielectric:

$$C = \frac{Q}{U} = \varepsilon_0 \varepsilon_r \frac{A}{d} \tag{9.79}$$

9.5.1.4 Energy Storage

Work must be done by an external influence to move charges between the conductors in a capacitor. When the external influence is removed, the charge separation persists and energy is stored in the electric field. If charge is later allowed to return to its equilibrium position, the energy is released. The work done in establishing the electric field, and hence the amount of energy stored, is given by equation (9.80) and for the parallel-plate capacitor by the use of equation (9.79) according to equation (9.81).

$$W_{el} = \frac{1}{2} C U^2 = \frac{1}{2} \frac{Q^2}{C} \tag{9.80}$$

$$W_{el} = \frac{1}{2}\varepsilon_0\varepsilon_r\frac{A}{d}U^2 \tag{9.81}$$

This stored electric energy can be used to perform mechanical work according to equation (9.82).

$$W_{mech} = Fx \tag{9.82}$$

9.5.2 Designs of Capacitive Actuators with Air-Gap

A preferred setup of electrostatic actuators is given by parallel-plate capacitors with air-gap. In these designs one electrode is fixed to the frame, while the other one is attached to an elastic structure, enabling the almost free movement in the intended direction (DOF). All other directions are designed stiff enough to prevent a significant displacement of this electrode. To perform physical work (displacement of the plate) the energy of the electric field according to equation (9.81) is used. Considering the principle design of these actuators two basic variants can be distinguished: the displacement may result in a change of the distance d, or the overlapping area A. Both variants are subject of discussion in the following paragraphs.

9.5.2.1 Movement Along Electric Field

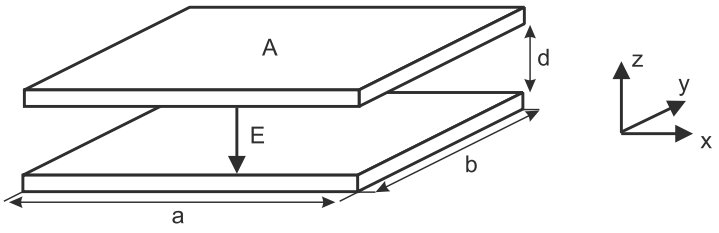


Fig. 9.67 Parallel-plate capacitor with air-gap

Looking at the parallel-plate capacitor from figure 9.67, the capacity C_L can be calculated with

$$C_L = \varepsilon_0 \cdot \frac{A}{d} \tag{9.83}$$

As shown before the stored energy W_{el} can be calculated for an applied voltage U:

$$W_{el} = \frac{1}{2}CU^2 = \frac{1}{2}\varepsilon_0 \frac{A}{d}U^2 \tag{9.84}$$

The force between both plates in z-direction can be derived by the principle of virtual displacement:

$$F_{z,el} = \frac{\partial W}{\partial z} = \frac{1}{2}U^2 \frac{\partial C}{\partial z} \tag{9.85}$$

$$\mathbf{F}_{z,el} = -\varepsilon_0 \frac{A}{2d^2}U^2 \mathbf{e}_z \tag{9.86}$$

The inhomogeneities of the electric field at the borders of the plates are neglected for this calculation, which is an acceptable approximation for the given geometrical relations of a large plate surface A and a comparably small plate distance d. A spring pulls the moving electrode into its idle position. Consequently, the actuator has to work against this spring. The schematic sketch of this actuator is shown in figure 9.68. The plate distance d is limited by the thickness of the insulation layer d_I. Analyzing the balance of forces according to equation (9.87) the interdependency of displacement z and electrical voltage U can be calculated:

$$F_z(z) = F_{spring}(z) + F_{z,el}(U,z) = 0 \tag{9.87}$$

$$-k \cdot z - \frac{1}{2}\varepsilon_0 A \frac{U^2}{(d+z)^2} = 0 \tag{9.88}$$

$$U^2 = -2\frac{k}{\varepsilon_0 A}(d+z)^2 \cdot z \tag{9.89}$$

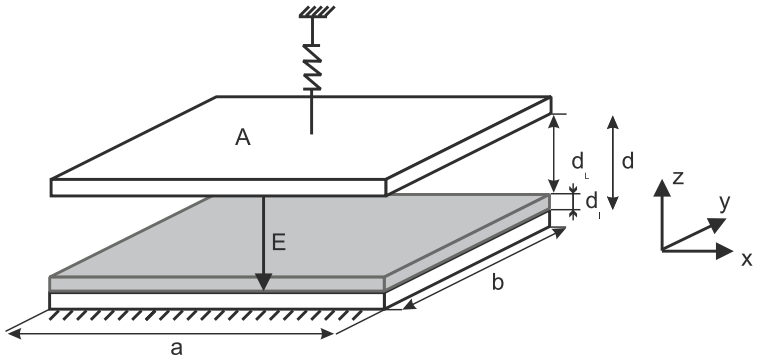


Fig. 9.68 Schematic setup of an actuator with variable air-gap

Analyzing the electrical voltage U in dependancy of the displacement z, a maximum can be identified:

$$\frac{\mathrm{d}U^2}{\mathrm{d}z} = -2\frac{k}{\varepsilon_0 A}(d^2 + 4dz + 3z^2) = 0 \tag{9.90}$$

$$z^2 + \frac{4}{3}dz + \frac{1}{3}d^2 = 0$$

$$z_1 = -\frac{1}{3}d;\ z_2 = -d \tag{9.91}$$

To use the actuator in a stable state, the force of the retaining spring has to be larger than the attracting forces of the charged plates. This condition is fulfilled for distances z

$$0 > z > -\frac{1}{3}d$$

Smaller distances cause attracting forces larger than the retaining force and the moving plate is strongly pulled onto the fixed plate ("pull-in" effect). As this would immediately result in an electric short cut, typical designs include an insulating layer on at least one plate. Equations 9.89 and 9.91 are used to calculate the operating voltage for the pull-in:

$$U_{pull-in} = \sqrt{\frac{8}{27}\frac{k}{\varepsilon_0 A}d^3} \tag{9.92}$$

The retention force to keep this state is much less than the actual force at the point of pull-in. It should be noted that force increases quadratically with decreasing distance. A boundary value analysis for $d \to 0$ provides the force $F \to \infty$. Consequently, the insulation layer fulfills the purpose of a force limitation, too.

9.5.2.2 Moving Wedge Actuator

A special design of air-gap actuators with varying plate distance is given by the moving wedge actuator. To increase displacement, a bended flexible counter-electrode is placed on a base electrode with a non-conductive layer. The distance between the electrodes increases wedge-like from the fixation to its free end. The resulting electrical field is higher inside the area where the flexible electrode is closest to the counter-electrode and it decreases with increasing air-gap. When designing the stiffness of the flexible electrode it has to be guaranteed, that it is able to roll along with the tightest wedge on the isolation. Figure 9.69 shows the underlying principle in idle-state and during operation [216].

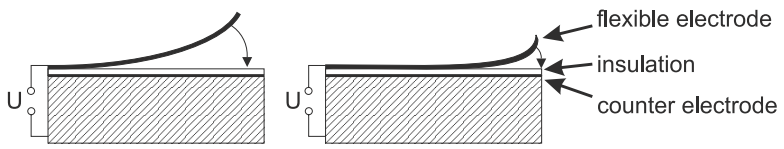


Fig. 9.69 Schematic view of a moving wedge actuator

9.5.2.3 Movement Perpendicular to Electric Field

The major difference compared to the prior design is given by the fact, that the plates are moving in parallel to each other. The plate distance d is kept constant, whereas the overlapping area varies. Analog to equation (9.84) the forces for the displacement can be calculated in both direction of the plane:

$$F_x = \frac{\partial W}{\partial x} = \frac{1}{2} U^2 \frac{\partial C}{\partial x} \tag{9.93}$$

$$\mathbf{F}_x = \frac{1}{2} \varepsilon_0 \frac{b}{d} U^2 \mathbf{e}_x \tag{9.94}$$

$$F_y = \frac{\partial W}{\partial y} = \frac{1}{2} U^2 \frac{\partial C}{\partial y} \tag{9.95}$$

$$\mathbf{F}_y = \frac{1}{2} \varepsilon_0 \frac{a}{d} U^2 \mathbf{e}_y \tag{9.96}$$

The forces are independent on the overlapping length only. As a consequence they are constant for each actuator position. Figure 9.70 shows the moving electrode being attached to a retaining spring.

If an electrical voltage is applied on the capacitor, the surface A increases along the border a. Hence, the spring is deflected and generates a counter force $\mathbf{F}_F$ according to

$$\mathbf{F}_F = -kx\mathbf{E}_x \tag{9.97}$$

The equilibrium of forces acting upon the electrode is given by

$$F_x(x) = F_F(x) + F_{x,el}(U) \tag{9.98}$$

From idle position ($F_x(x) = 0$) the displacement of the electrode in x-direction is given by:

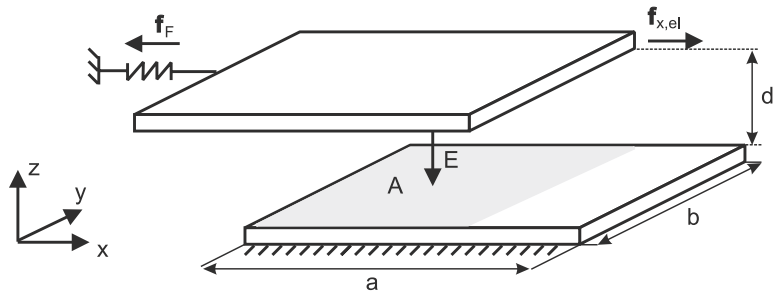


Fig. 9.70 Electrostatic actuator with variable overlapping area

$$x = \frac{1}{2}\varepsilon_0 U^2 \frac{b}{d}\frac{1}{k} \tag{9.99}$$

Typically, this design is realized in a comb-like structure, with one comb of electrodes being engaged in a counter electrode comb. This equals an electrical parallel circuit of n capacitors, which is identical to a parallel circuit of force sources complementing each other. Figure 9.71 shows such a design. The area of the overlapping electrodes is given by a in x-direction and b in y-direction. With the plate distance d the capacity according to equation (9.100) can be calculated.

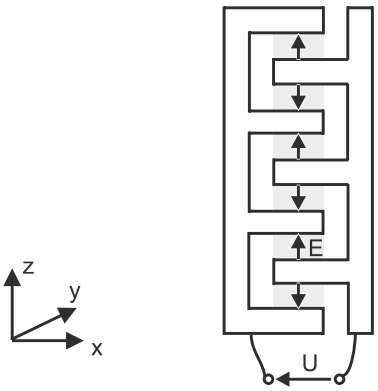


Fig. 9.71 Actuator with comb-electrodes and variable overlapping area

$$C_Q = \varepsilon_0 \cdot \frac{ab}{d} \cdot n \tag{9.100}$$

By differentiating the energy according to the movement direction the electromotive force can be calculated:

$$F_x = \frac{\partial W}{\partial x} = \frac{1}{2}U^2 \frac{\partial C}{\partial x} = \frac{1}{2}U^2 \varepsilon_0 \frac{b}{d} \cdot n \tag{9.101}$$

9.5.2.4 Summary and Examples

For all actuators shown, the electrostatic force acts indirect against the user and is transmitted by a moveable counter-electrode. A much simpler design of tactile displays makes use of the user's skin as a counter electrode, which experiences the whole electrostatic field force. Accordingly, tactile electrostatic applications can be distinguished in direct and indirect principles.

Direct Field Force

The simplest design combines one electrode, respectively a structured electrode array, and an isolating layer. A schematic sketch is given in figure 9.72. The user and his finger resemble the counter electrode. With the attractive force between the conductive skin and the electrodes, a locally distributed increase of friction can be achieved. It hinders a relative movement and can be perceived by the user. Such systems can be easily realized and excellently miniaturized. Their biggest disadvantage is their sensitivity on humidity on the surface, which is brought onto the electrodes during any use in form of sweat. This leads to a blocking of the electrical by the conductive sweat layer above the isolation, preventing the user to feel any relevant force.

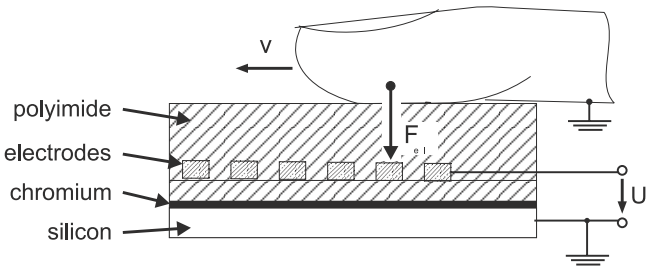


Fig. 9.72 Electrostatic stimulator with human finger as coutner electrode [253].

Indirect Field Force

With these systems the field force is used to move an interacting surface. The finger of the user interacts with these surfaces (sliders) and experiences their movements

as a perceivable stimulation. A realization with a comb of actuators moving orthogonal to the field direction is given in figure 9.73. The structural height is 300 μm providing 1 mN at operating voltages of up to 100 V. The same design with an actuator made of parallel electrodes can achieve displacements of 60 μm. The comb-electrodes shown here displace 100 μm.

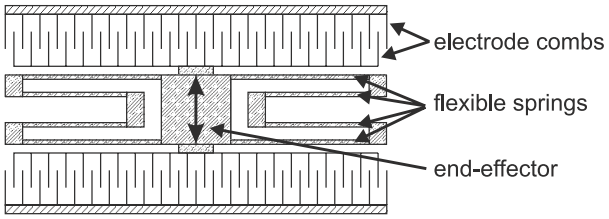


Fig. 9.73 Electrostatic comb-actuator for tangential stimulation [71].

Summary

Electrostatic drives with air gap achieve force in the range of mN to N. As the actuators are driven by fields, the compromise between plate distance and electrical operation voltage has to be validated for each individual application. The breakdown field strength of air (approx. 3V/μm) is the upper, limiting factor. The actuators' displacement is limited to several μm. At the same time the operating voltages reach several hundred volts. Due to the very low achievable displacement the application of electrostatic actuators is limited to tactile stimulations only. For the concrete actuator design it is recommended to deal with the modeling of such actuators, e.g. based on the theory of concentrated networks according to LENK [158]. This allows the analysis of the complete electromechanical system starting from the applicable mechanical load situation to the electrical control with a single methodological approach.

9.5.3 Dielectric Elastomer Actuators

As with many other areas, new synthetic materials replace classic materials such as metals in actuator design. Thanks to the enormous progress in material development their mechanical properties can be adjusted to a large spectrum of possible applications. Other big advantages are given by the very cheap material costs. Additionally, almost any geometrical shape can be manufactured with relatively small efforts.

Polymers are called “active polymers” if they are able to change their shape and form under the influence of external parameters. The causes for theses changes may be manifold: electric and magnetic fields, light and even ph-value. Being used within actuators, their resulting mechanical properties like elasticity, applyable force and deformation at simultaneously high toughness and robustness are quite comparable to biological muscles [10].

To classify the large variety of “active polymers” they are usually distinguished according to their physical working principle. A classification into “non-conductive polymers”, activated e.g. by light, pH-value or temperature, and “electrical polymers”, activated by an arbitrary electrical source. Latter are called “electroactive polymers” (EAP) and are further distinguished into “ionic” and “electronic” EAPs. Generally speaking electronic EAP are operated at preferably high field strengths near the breakdown field strength. Depending on the layer thickness of the dielectrics $1-20\,\mathrm{kV}$ are typical operation voltages. Consequently, very high energy densities at low reaction times (in the area of milliseconds) can be achieved. In contrast, ionic EAP are operated at obviously lower voltages of $1-5\,\mathrm{V}$. However, an electrolyte is necessary for transportation of the ions. It is frequently provided by a liquid solution. Such actuators are typically realized as bending bars, achieving large deformations at their tip with long reaction times (several seconds).

All EAP technologies are subject of actual research and fundamental development. However, two actuator types are already used in robotics: “Ionic polymer metal composite” (IPMC) and “dielectric elastomer actuators” (DEA). A summary and description of all EAP-types is offered by KIM [137]. Their functionality will be further discussed in the following paragraphs as they affiliate to the group of electrostatic actuators. A comparison between characteristic values of dielectric elastomer actuators and the muscles of the human is shown in table 9.7. By the use of an elastomer actuator with large expansion additional mechanical components such as gears or bearings are needless. Additionally, the use of these materials may be combined to complex designs similar to and inspired by nature. One application is e.g. the locomotion of insects and fishes within bionic research [11].

Table 9.7 Comparison of human muscle and DEA according to Pei [196]

parameter	human muscle	DEA
strain (%)	$20-40$	10 bis >100
stress (MPa)	$0,1-0,35$	$0,1-2$
energy density (kJ/m^3)	$8-40$	$10-150$
density (kg/m^3)	1037	≈ 1000
velocity of deformation (%/s)	>50	450 (acrylic)
		34000 (silicone)

9.5.3.1 Dielectric Elastomer Actuators - Electrostatic Solid State Actuators

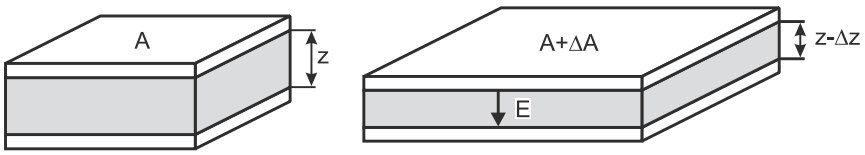


Fig. 9.74 DEA in initial state (left) and charged state (right).

The design of dielectric elastomer actuators is identical to the design of a parallel-plate capacitor, but with an elastic dielectric (a polymer respectively elastomer) sandwiched by compliant electrodes. Hence, it is a solid state actuator. The schematic design of a dielectric elastomer actuator is visualized in figure 9.74, left. In an uncharged condition the capacity and the energy stored is identical to an air-gap actuator (equations 9.79 and 9.80). A change of this condition happens by the application of a voltage U and is visualized in figure 9.74, on the right: The charged capacitor contains more charges ($Q + \Delta Q$), the electrode area increases ($A + \Delta A$), whereas the distance ($z - \Delta z$) simultaneously decreases. The change of energy after an infinitesimal change dQ, dA and dz is calculated in equation 9.102:

$$\mathrm{d}W = \left(\frac{Q}{C}\right)\mathrm{d}Q + \left(\frac{1}{2}\frac{Q^2}{C}\frac{1}{z}\right)\mathrm{d}z - \left(\frac{1}{2}\frac{Q^2}{C}\frac{1}{A}\right)\mathrm{d}A \tag{9.102}$$

$$\mathrm{d}W = U\mathrm{d}Q + W\left[\left(\frac{1}{z}\right)\mathrm{d}z - \left(\frac{1}{A}\right)\mathrm{d}A\right] \tag{9.103}$$

The internal energy change equals the change of the electrical energy by the voltage source and the mechanical energy used. Latter depends on the geometry (parallel (dz) and normal (dA) to the field's direction). In comparison to the air-gap actuator in section 9.5.2 an overlay of decreasing distance and increasing electrodes' area occurs. This is causes by a material property which is common to all elastomers and to almost all polymers: the aspect of volume constancy. A body compressed in one direction will extend in the remaining two other dimensions if it is incompressible. This gives a direct relation between distance change and the change of electrodes' area. As a consequence equation (9.104) results

$$A\,\mathrm{d}z = -z\mathrm{d}A \tag{9.104}$$

simplifying equation (9.103) to

$$\mathrm{d}W = U\mathrm{d}Q + 2W\left(\frac{1}{z}\right)\mathrm{d}z \tag{9.105}$$

The resulting attractive force of the electrodes can be derived from this electrical energy. With respect to the electrode surface A the electrostatic pressure p_{el} at $\mathrm{d}Q = 0$ is given according to equation (9.106)

$$p_{el} = \frac{1}{A}\frac{\mathrm{d}W}{\mathrm{d}z} = 2W\frac{1}{Az} \tag{9.106}$$

and by the application of equation (9.80)

$$p_{el} = 2\left(\frac{1}{2}\varepsilon_0\varepsilon_r Az\frac{U^2}{z^2}\right)\frac{1}{Az} = \varepsilon_0\varepsilon_r E^2 \tag{9.107}$$

Comparing this result with equation (9.85) as a reference for a pressure of an air-gap actuator with variable plate distance, dielectric elastomer actuators are capable of generating a pressure twice as high with otherwise identical parameters [198].

Additional reasons for the obviously increased performance of the dielectric elastomer actuators are based on their material, too. The relative permittivity is given by $\varepsilon_r > 1$, depending on the material $\varepsilon_r \simeq 3 - 10$. By chemical processing and implementation of fillers the relative permittivity may be increased. However, it has to be noticed that other parameters (such as the breakdown field strength and the e-modulus) may become worse, possibly the positive effect of the increased ε_r gets lost. Especially the breakdown field strength is one of the most limiting factors. With many materials an increase in breakdown field strength could be observed after planar prestrain. In these cases breakdown field strengths of $100 - 400\,\mathrm{V}/\mu\mathrm{m}$ are typical.

The pull-in effect does not happen at $z = 1/3 \cdot z_0$ (air-gap actuators), but at much higher deflections. With some materials mechanical prestrain of the actuator allows to displace the pull-in further, reaching the breakdown field strength before. The reason for this surprising property is the volume constant dielectric layer showing viscoelastic properties. It complies with a return spring with strong nonlinear force-displacement characteristics for large extensions. Its working point is displaced along the stress-strain-curve of the material as an effect of the mechanical prestrain.

For the application in dielectric elastomer actuators many materials may be used. The material properties cover an extreme wide spectrum ranging from gel-like polymers up to relatively rigid thermoplastics. Generally speaking, every dielectric material has to have a high breakdown field strength and elasticity beside a high relative permittivity. Silicone provides highest deformation-velocities and a high temperature resistance. Acrylics have high breakdown field strength and achieve higher energy densities. The following list is a selection of the dielectric materials most frequently used today:

- silicone
  HS 3 (Dow Corning)
  CF 19-2186 (Nusil)
  Elastosil P7670 (Wacker)
- acrylics

VHB 4910 (3M)

The most frequently used materials for the elastic electrodes are graphite powder, conductive carbon, and carbon grease.

9.5.4 Designs of Dielectric Elastomer Actuators

As mentioned before, dielectric elastomer actuators achieve high deformations (compression in field direction) of $10-30\%$. To keep voltages within reasonable ranges, layer thicknesses of $10-100\,\mu$m are used depending on the breakdown field strength. The resulting absolute displacement in field direction is too low to be useful. Consequently, there are several concepts to increase it. Two principle movement directions are distinguished for this purpose: the longitudinal effect in parallel to the field (thickness change), and the transversal effect orthogonal to field (surface area change). The importance of this discrimination lies in the volume constancy of the material: uni-axial pressure load equals a two-axial tension load in the remaining spatial directions. Hence, two transversal tensions within the surface result in a surface change. For materials fulfilling the concept of "volume constancy", equation (9.108) is valid, providing the following properties for the longitudinal compression S_z and the transversal elongation S_x:

- with a longitudinal compression of 62 % both extensions are identical;
- for smaller values of the longitudinal compression the resulting transversal extension is smaller;
- for a longitudinal compression $> 62\%$ of the transversal extensions increases much faster than the longitudinal compression.

$$S_x = \frac{1}{\sqrt{1-S_z}} - 1 \tag{9.108}$$

The extension of the surface area S_A depends on the longitudinal compression S_z according to equation (9.109):

$$S_A = \frac{dA}{A} = \frac{S_z}{1-S_z} \tag{9.109}$$

The increase of the area with uni-axial compression is always larger than the change of thickness. Actuators built according to this principle are the most effective ones. Figure 9.75 shows three typical designs. A roll-actuator (left) being built as full or tubular cylinder can achieve length changes of more than 10 %. KORNBLUH [143] describes an acrylic roll-actuator achieving a maximum force of 29 N with an own weight of no more than $2,6$ g at a extension of 35 mm. The manufacture of electrodes with a large area is very simple. On the other hand the rolling of the actuators with simultaneous prestrain (up to 500 %) can be very challenging. With a

stack-actuator (middle) very thin dielectric layers with minimized operational voltages can be achieved, depending mainly on the manufacturing technique. As the longitudinal effect is used extension are limited to approximately 10%. However, due to their design and fabrication process actuator arrays at high density can be built [119]. The most simple and most effective designs are based on a restrained foil, whose complete surface change is transformed into an uparching (diaphragm-actuator, right) [144]. If this actuator experiences a higher external load, such as from a finger, an additional force source, e.g. a pressure, has to be provided to support the actuators own tension.

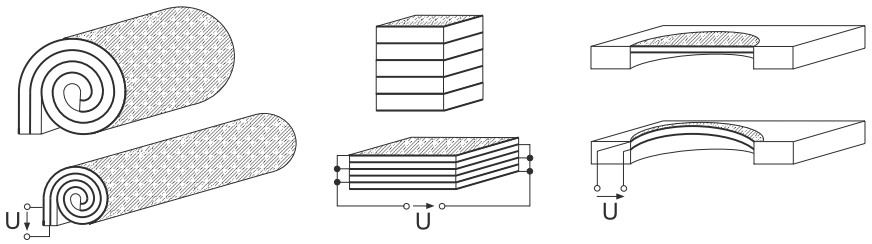


Fig. 9.75 Typical designs of Dielectric Elastomer Actuators: roll-actuator (left), stack-actuator (center) and diaphrag-actuator (right).

9.5.4.1 Summary and Examples

As with air-gap actuators, a dielectric solid state actuator's major limit is given by the breakdown field strength of the dielectric. However, in contrast to the air-gap actuators, a carefully chosen design can easily avoid any pull-in effect. Consequently, these actuators show a larger workspace; and with the high number of different design variants a wide variety of applications can be realized, just depending on the requirements on displacement, maximum force and actuator density.

Tactile Displays

The simplest application of a tactile display is a Braille-device. Such devices are meant to display Braille-letters in patterns of small, embossed dots. In standard Braille six dots are being used, in computer compatible Euro-Braille eight dots. These points are arranged in a 2x3 or 2x4 matrix, respectively (9.76). In a display-device 40 to 80 characters are displayed simultaneously. In state-of-the-art designs each dot is actuated with one piezoelectric-bending actuator (section 9.4.5). This technical effort is reason for the high price of these devices. As a consequence there

are several functional samples existing, which prove the applicability of less expensive drives with simplified mechanical designs, but still sufficient performance. Each of the three variants for dielectric elastomer actuators has already been used for this application.

Figure 9.76 shows the schematic sketch of roll-actuators formed to an actuator column [206]. Each roll-actuator moves one pin, which itself is pushed up above the base plate after applying a voltage. The elastomer film is coiled around a spring of 60 mm length with a diameter of 1.37 mm. With an electric field of $100\,\mathrm{V}/\mu\mathrm{m}$ applied, the pretensioned spring can achieve a displacement of 1 mm at a force of 750 mN. The underlying force source is the spring with spring-constant 225 N/m pretensioned by a passive film. The maximum necessary displacement of $500\,\mu\mathrm{m}$ is achieved at field strengths of $60\,\mathrm{V}/\mu\mathrm{m}$.

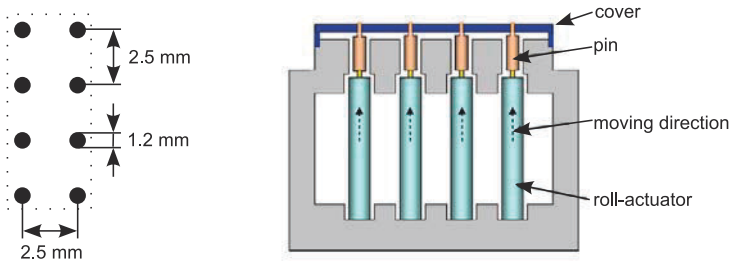


Fig. 9.76 Presenting a Braille sign with roll-actuators. left: geometry ; right: schematic setup of a Braille row [206].

The application of stack actuators according to JUNGMANN [119] is schematically sketched in figure 9.77, left. The biggest advantage of this variant is given by the extremely high actuator density combined with a simple manufacturing process. Additionally, the closed silicone elements are flexible enough to be mounted on an almost arbitrary formed surface. The surface - in itself being made of silicone - shows an adequate roughness and temperature conductivity. It is perceived as "convenient" by many users. With a field strength of $30\,\mathrm{V}/\mu\mathrm{m}$ a stack made of 100 dielectric layers displacements of $500\,\mu\mathrm{m}$ can be achieved. The load of a reading finger on the soft substrate generates a typical contact pressure of 4 kPa resulting in a displacement of $25\,\mu\mathrm{m}$. This extension is considerably less than the perception threshold of 10 % of the maximum displacement. For control of the array it has to be noted, that the actuators are displaced in a negative logic. With applied voltage the individual pin is pulled downwards.

The design of a Braille display with diaphragm-actuators according to HEYDT [94] demonstrates the distinct properties of this variant. The increase of elastomer surface results in a notable displacement of a pin from the device's contact area. However, a mechanical prestrain is necessary to provide a force. This can be either generated by a spring or air pressure below the actuator. Figure 9.77 on the right

gives a schematic sketch for a single point being pretensioned by a spring with a diameter of 1.6 mm. At an operating voltage of 5.68 kV the actuator displaces in idle mode 450 μm.

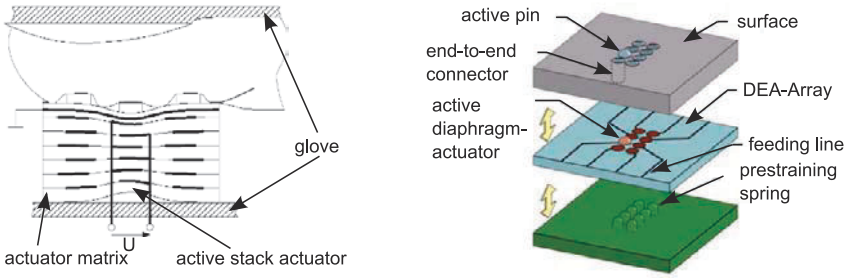


Fig. 9.77 left: Actuator row with stack-actuators [119]; right: Use of diaphrag-Actuators [94].

Artificial Muscles

To advance research and development and to market the future potential of adaptive material-systems media-friendly, YOSEPH BAR-COHEN (Jet Propulsion Laboratorys of the NASA at "Institute of Technology") initiated a competition in California in 2001. In March 2005 during the SPIE's Smart Structures and Materials Symposium (Conference on Electriactive Polymer Actuators and Devices EAPAD) the first contest between a 17 years old girl and several machines actuated by EAP actuators took place. With two actuator variants from DEA and IPMC a synthetic arm from Environmental Robots Incorporated (ERI) was able to generate sufficient counter forces for 26 seconds, and achieved a stalemate situation. The second rank was achieved by the arm of the Suisse "Eidgenössische Material- und Prüfanstalt (EMPA)". The used "muscles" have been made of seven roll-actuators and have lifted a mass of approximately 38 kg.

9.5.5 Electro-Rheological Fluids

Fluids being influenced in their rheological properties (especially the viscosity) by electrical field varying in direction and strength are called electro-rheological fluids (ERF). Consequently ERF are classified as non-Newton fluids, as they have a variable viscosity at constant temperature. The electro-rheological effect has been observed for the first time in 1947 at a suspension of cornstarch and oil by WILLIS WINSLOW. Electro-rheological fluids include dipoles made of polarized particles,

which are dispersed in a conductive suspension. These particles are aligned in an applied electrical field. An interaction between particles and free charge carriers happens. Chain-like microstructures are built between the electrodes [41, 100, 227] in this process. However, it seems as if this is not the only effect responsible for the viscosity change, as even when the microstructures [194] were destroyed a significant viscosity increase remained. The exact analysis of the mechanism responsible for this effect is subject of actual research. The viscosity of the fluid changes depending on the strength of the applied electrical field. With an electric field of $1-10\,\mathrm{kV/mm}$ the viscosity may change up to a factor of 1000 compared to the field-free state. This enormous change equals a viscosity difference between water and honey. A big advantage of this method can be found in the dynamics of the viscosity change. It is reversible and can be switched within one millisecond. Therefore, electro-rheological fluids are suitable for dynamic applications, too. If large field strengths are assumed, the ERF can be modelled as BINGHAM-fluid. It has a threshold for linear flow characteristics: starting at a minimum tension $\tau_{F,d}$ (flow threshold) the fluid actually starts to flow. The fluid starts flowing right below this threshold. The shear forces τ are calculated according to equation (9.110):

$$\tau = \mu\dot{\gamma} + \tau_{F,d} \tag{9.110}$$

With μ being the dynamic viscosity, $\dot{\gamma}$ the shear rate, and $\tau_{F,d}$ is the dynamic flow limit. Latter changes quadratically with the electrical field strength (equation (9.111)). The proportional factor C_d is a constant provided with the material's specifications.

$$\tau_{F,d} = C_d E^2 \tag{9.111}$$

For complex calculations modelling the fluid's transition to and from the state of flow, the model is extended to a nonlinear system according to equation (9.112) (for $n = 1$ this equals equation (9.110))

$$\tau = \tau_{F,d} + k\dot{\gamma}^n \tag{9.112}$$

This general form describes the shear force for visco-plastic fluids with flow-limit according to VITRANI [273]. For an analysis of idle state with shear rate $\dot{\gamma} = 0$ the static flow-limit $\tau_{F,s}$ with $\tau_{F,s} > \tau_{F,d}$ is introduced. When exceeding the static flow limit, the idle fluid is deformed. With the specific material constants C_s and E_{ref} equation (9.113) can be formulated:

$$\tau_{F,s} = C_s(E - E_{ref}) \tag{9.113}$$

The materials used for the particles are frequently metal oxides, silicon anhydride, poly urethane and polymers with metallic ions. The diameter of particles is $1-100\,\mu m$, their proportion on the fluid's volume is 30 to 50 %. As carrier medium typically oils (such as silicone oil) or specially treated hydrocarbon are used. To additionally improve the viscosity change, nanoscale particles are added in the electro-rheological fluids, too ("giant electro-rheological effect" [75], [283]). In [49] and

[202] further mathematical modelling is presented for the dynamic flow behaviour of ER-fluids. The central property of ERF - to reversibly change the viscosity - is used for force-feedback devices, haptic displays, and artificial muscles and joints. As the change in viscosity is mainly a change in counter-forces but not in shape or direct-forces, ERF actuators are counted to the group of "passive actuators". For the characterization of their performance, the ratio between stimulated and idle state is used. They are built in three principle design variants [33]:

9.5.5.1 Shear Mode

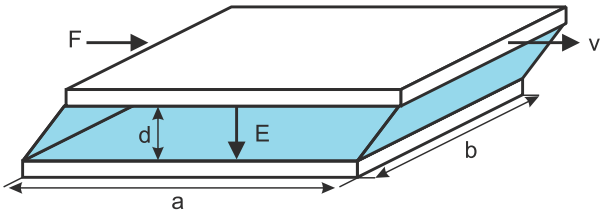


Fig. 9.78 Using ERF to vary the shear force.

The ER-fluid is located in between two parallel plates, one fixed and one moving relatively to the fixed one. The only constrain is given by a fixed inter-plate distance d. If a force F is applied on the upper plate, it is displaced by a value x at a certain velocity v. For the configuration shown in figure 9.78 the mechanical control ratio λ can be calculated according to equation (9.116) from the ratio of dissipative forces (field-dependent flow-stresses, equ. 9.115) and the field independent viscosity-term (equation (9.114)) [199]. η gives the basis viscosity of the ER-Fluid (in idle state) and τ_y the low-stress depending on the electrostatic field.

$$F_\eta = \frac{\eta v a b}{d} \tag{9.114}$$

$$F_\tau = \tau_y a b \tag{9.115}$$

$$\lambda = \frac{F_\tau}{F_\eta} = \frac{\tau_y d}{\eta v} \tag{9.116}$$

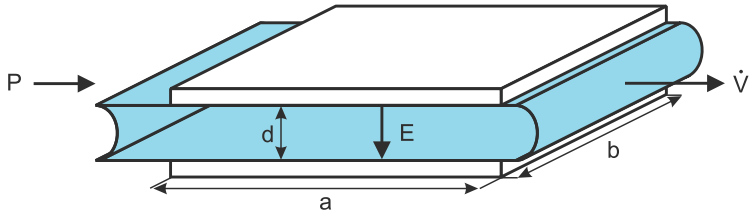


Fig. 9.79 Varying the flow channel's resistivity with ERF-actuators

9.5.5.2 Flow Mode

The schematic sketch of this configuration is shown in figure 9.79. Both fixed plates form a channel, with the fluid flowing through it due to an external pressure difference p and a volume-flow $\dot{V}$. With an electric field E applied between the plates the pressure loss increases along the channel and the volume flow is reduced. Analog to the prior design a field independent viscosity based pressure loss p_η and a field dependent pressure loss p_τ can be calculated [199]:

$$p_\eta = \frac{12\eta\dot{V}a}{d^3 b} \tag{9.117}$$

$$p_\tau = \frac{c\tau_y a}{d} \tag{9.118}$$

The mechanical control ratio equals

$$\lambda = \frac{p_\tau}{p_\eta} = \frac{c\tau_y d^2 b}{12\eta\dot{V}} \tag{9.119}$$

At an adequate dimensioning of the fluid, the flow-resistance can be increased by the electrical field to such a degree, that the fluid stops completely when exceeding a specific voltage. This makes the channel a valve without any moving mechanical components.

9.5.5.3 Squeeze Mode

A design to generate pressure is schematically sketched in figure 9.80. In contrast to the variants shown before, the distance between both plates is subject to change now. If a force acts on the upper plate, it moves downwards. This results in the fluid being pressed outside. A plate distance d_0 is assumed at the beginning, and a relative movement of v of the plate moving downwards. The velocity dependent viscosity force F_η and the field dependent tension term F_τ [117] are calculated according to:

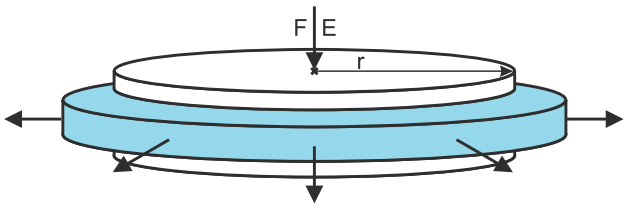


Fig. 9.80 Varying the acoustic impedance with ERF-actuators under external forces

$$F_\eta = \frac{3\pi\eta v r^4}{2(d_0 - z)} \tag{9.120}$$

$$F_\tau = \frac{4\pi\tau_y r^3}{3(d_0 - z)} \tag{9.121}$$

which gives the mechanical control ratio:

$$\lambda = \frac{8\tau_y 3(d_0 - z)^2}{9\eta v r} \tag{9.122}$$

With pressure (force on the upper plate) the fluid is pressed out of the gap. In this configuration the force-displacement-characteristics is strongly influenced by the electrical field strength. An analysis of the dynamic behaviour of such an actuator is described in [277].

9.5.5.4 Designing ERF-Actuators

The maximum force F_τ and the necessary mechanical power P_{mech} are the input values for the design of ERF-actuators from the perspective of an application engineer. Equations 9.114 and 9.122 can be combined to calculate the necessary volume for providing a certain power with all three actuator configurations.

$$V = k\frac{\eta}{\tau_y^2}\lambda P_{mech} \tag{9.123}$$

Consequently, the volume is defined by the mechanical control ratio, the fluid-specific values η and τ_y, such as a constant k dependent on the actual configuration. The electrical energy W_{el} necessary to generate the electrostatic field of the actuator (volume-dependent) is calculated according to equation (9.124).

$$W_{el} = V(\frac{1}{2}\varepsilon_0\varepsilon_r E^2) \tag{9.124}$$

9.5.5.5 Comparison to Magneto-Rheological Fluids

Magneto-rheological fluids (MRF) are very similar to electro-rheological fluids. However, the physical properties of the fluids are influenced by magnetic fields. All calculations which are shown before are applicable to MRF, too. Looking at the volume necessary for an actuator according to equation (9.123), considering the viscosities of electro-rheological and magneto-rheological fluids being comparable, a volume ratio proportional to the reciprocal ratio of the fluid-tensions' square according to equation (9.125) results:

$$\frac{V_{ERF}}{V_{MRF}} = \frac{\tau_{MRF}^2}{\tau_{ERF}^2} \tag{9.125}$$

In a rough but good approximation the flow-stress of a magneto-rheological fluid is one magnitude larger than of an ERF, resulting in a smaller (approximately factor 100) volume of a MRF actuator compared to the ERF. However, a comparison between both fluids going beyond the pure volume analysis for similar output power is hard: for an ERF high voltages at relatively small currents are required. The main power leakage is lost by leakage-currents through the medium (ERF) itself. With MRF-actuators smaller electrical voltages at very high currents become necessary to generate an adequate magnetic field. The energy for a MRF-actuator is calculated according to equation (9.126) with the magnetic flux density B and the magnetic field strength H.

$$W_{el,MRF} = V_{MRF}(\frac{1}{2}BH) \tag{9.126}$$

The ratio between the energies for both fluids is calculated according to equation (9.127)

$$\frac{W_{el,ERF}}{W_{el,MRF}} = \frac{V_{ERF}}{V_{MRF}} \frac{\varepsilon_0 \varepsilon_r E^2}{BH} \tag{9.127}$$

With typical values for all parameters the necessary electrical energy for actuator control is comparable for both fluids. A good overview on the design of actuators for both types of fluids is given in [36].

9.5.5.6 Summary and Examples

Electro-rheological fluids are also called partly-active actuators, as they are not transforming the electrical values into a direct movement, but change their properties due to the electrical energy provided. The change of their properties covers a wide range. Naturally, their application in haptics ranges from small tactile displays to larger haptic systems.

Tactile Systems

The first application of ERF as tactile sensor in an artificial robot hand has been made in 1989 by KENALEY [127]. Starting from this work several ideas developed to use ERF in tactile arrays for improving systems for virtual reality applications. Several tactile displays, among them a 5x5 Matrix from TAYLOR [254] and another one from BÖSE [32], were built. Figure 9.81 shows the schematic design of such a tactile element. A piston is pressed in an ERF filled chamber by the user. Varying counter forces are generated depending on the actuation state of the ERF. Elastic foam is connected to the piston as a spring to move it back to its resting position. With an electric field of 3 V/μm a force of 3.3 N can be achieved at a displacement of 30 mm. Switching the electrical voltages is realized by light emitting diodes and corresponding receivers (GaAs-elements) on the backplane.

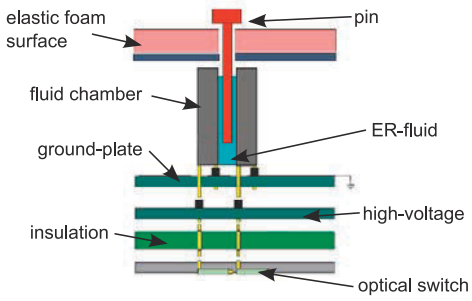


Fig. 9.81 Schematic setup of a tactile actuator [32]

Haptic Rotary Knob

Another obvious application for ERF in haptic systems is their use as a "variable brake". This is supported by the fact that typical applications beside haptic systems are variable brakes and bearings (e.g. adaptive dampers). There are several designs with a rotary knob mobbing a spinning disk within an ERF or MRF [162] generating varying counter torques. Figure 9.82 shows the schematics of a MRF-rotary knob. In this case the measurement of the rotary angle is solved by a potentiometer. In dependency on the rotary angle the intended counter force respectively torque is generated. The user can perceive a "latching" of the rotary knob with a mature system. The latching depth itself can be varied in a wide range. By the varying friction hard stops can be simulated, too, such as sticking and of course free rotation.

An extension of the one-dimensional system is presented in [281]. Two systems based on ERF are coupled to a joystick with two DOF. A counter force can be generated in each movement direction of the joystick. As ERF are able to generate

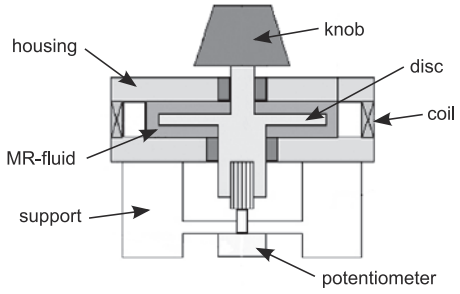


Fig. 9.82 Schematic setup of a haptic rotary knob [162]

higher torques with less energy required compared to a normal electrical drive, they are especially suitable to mobile applications like in cars.

Force-Feedback Glove

A force-feedback glove was designed as a component for a simulator of surgeries [12]. Surgical interventions shall be trained by the aid of haptic feedback. The system MEMICO ("Remote Mechanical Mirroring using Controlled stiffness and Actuators") shall enable a surgeon to perform the treatment with a robot in telemanipulation, whereas the haptic perception is retained. ERF actuators are used for both ends: on the side of the end-effector, and for the haptic feedback to the user. The adjustable elasticity is based on the same principle as with tactile systems. For generating forces a force source is necessary. A new ECFS actuator ("Electronic Controlled Force and Stiffness") is used for this application. The schematic design is shown in figure 9.83. It is an actuator according to the inchworm-principle, wherein both brakes are realized by the ER-fluid surrounding it. The driving component for the forward- and backward movement is realized by two electromagnets.

Both actuators are assembled within a haptic exoskeleton. They are mounted on the rim of a glove to conserve the mobility of the hand. With the actuators in between all finger-joints arbitrary forces and varying elasticities can be simulated independently. The ECFS actuators are operated at voltages of 2 kV and generate a force of up to 50 N.

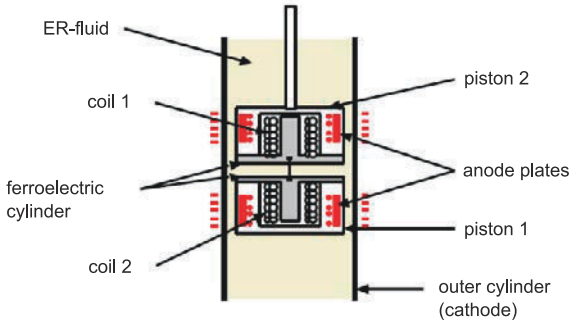


Fig. 9.83 Schematic setup of an ERF-Inchworm motor [12]

9.6 Special Designs of Haptic Actuators

THORSTEN A. KERN

The actuation principles discussed so far are the most common approaches to the actuation of haptic devices. Besides these principles, there are numerous research projects, singular assemblies, and special classes of devices. The knowledge of these designs is an enrichment for any engineer, yet it is impossible to completely cover the variety of all haptic designs in a single book. This section, nevertheless, intends to give a cross section of alternative, quaint and unconventional systems for generating kinaesthetic and tactile impressions. This cross-section is based on the authors' subjective observations and knowledge and does not claim to be exhaustive. The discussed systems have been selected, as examples suited best to cover one special class of systems and actuators, each. They are neither the first systems of their kind, nor necessarily the best ones. They are thought to be crystallization points of further research, if specific requirements have to be chosen for special solutions. The systems shown here are meant to be an inspiration and an encouragement not to discard creative engineering approaches to the generation of haptic impressions too early during the design process.

9.6.1 Haptic-Kinaesthetic Devices

Haptic-kinaesthetic devices of this category excel primarily due to their extraordinary kinematics and not to very special actuation principles. Nevertheless, every engineer may be encouraged to be aware of the examples of this device class and let this knowledge influence his / her own work.

9.6.1.1 Spidar-System

The Spidar (fig. 9.84) is based on the work of Prof. SATO and has frequently been used in research projects [247, 181] as well as in commercial systems. It is composed of an interaction handle - usually a ball - held by eight strings. Each string is operated by an actuator, which is frequently (but not obligatorily) mounted in the corners of a rectangular volume. The drives are able to generate pulling forces on the strings, enabling the generation of forces and torques in six DOFs on the handle. Typically the actuators used are based on electrodynamic electronic - commutated units. The Spidar-system can be scaled to almost any size, ranging from table-top devices to room-wide installations. It convinces by the small number of mechanical components and the very small friction. As strings are able to provide pull forces

only, it is worth noting that just two additional actuators are sufficient to compensate this disadvantage.

a) 

b) 

Fig. 9.84 a) Desktop-version of the Spidar with ball-like interaction handle, b) room-size version "INCA 6D" with 3D visualization environment by Haption.

9.6.1.2 MRI-Compatible Haptic Device

The design of electrodynamic and electromagnetic actuators always includes a magnetic field and the corresponding flux guiding materials. Yet, when designing actuators for medical applications to be used near or in a magnet-resonance tomography (MRT, MRI) the developer is limited in his/her choice of materials. In the presence of the strong static magnetic fields ferromagnetic materials are not applicable. At the same time conductive materials cannot be used due to the high frequencies of the switched gradient-fields inducing currents. For the application in haptic systems inside or near a MRT, in [208] an electromagnetic actuator (fig. 9.85) was developed and tested. It uses the existing static magnetic field of MRI as a counterpole to the variable magnetic field generated by two coils. The torques generated reach impressive 4 Nm at a distance of 1 m to the MRI, without influencing the imaging process significantly.

9.6.1.3 Magnetorheological Fluids as Three-Dimensional Haptic Display

The wish to generate an artificial haptic impression in a volume for free interaction is one of the major motivations for many developments. The rheological systems shown in section 9.5 provide one option to generate such an effect. For several years the team of BICCHI has been working on the generation of spatially resolved areas of differing viscosity in a volume (fig. 9.86) to generate force-feedback on an exploring hand. Lately, the results were summarized in [20]. The optimization of such actuators is largely dependent on the control of the rheological fluid [21]. The psycho-physical experiments performed until today show that the identification of

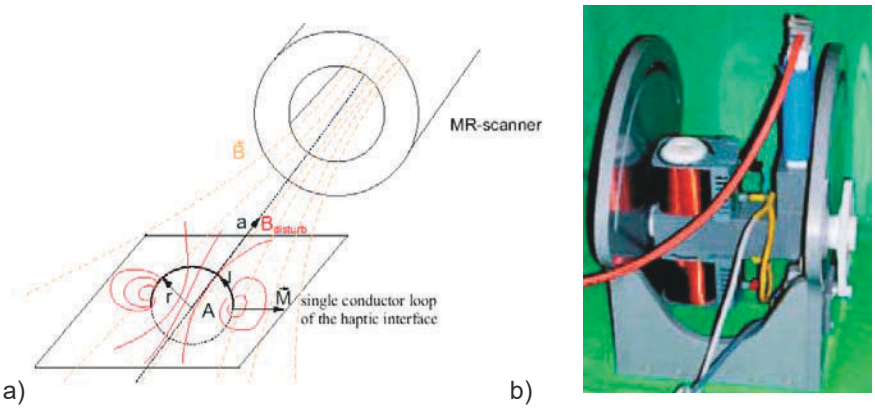


Fig. 9.85 Working-principle of the MRI actuator by [208] (a) and actual design with coils visible on its left side (b).

simple geometrical structures can be achieved on the basis of a 4x4 pattern inside the rheological volume.

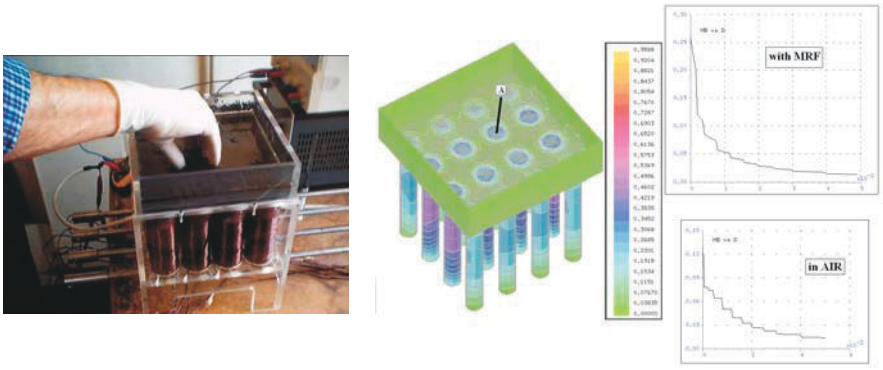


Fig. 9.86 Magnetorheological actuation principle for full-hand interaction based on a 4x4 pattern [20].

9.6.1.4 Self-Induction and Eddy Currents as Damping

An active haptic device is designed to generate forces resp. torques in any direction. By the concept of "active" actuation the whole spectrum of mechanical interaction objects (e.g. masses, springs, dampers, other force sources like muscles, and moving objects) is covered. Nevertheless, only a slight portion of haptic interaction actually

is “active”. This has the side effect (of control engineering approaches) that active systems have continuously to be monitored for passivity. An alternative approach to the design of haptic actuators is given by choosing technical solutions able to dissipate mechanical energy. A frictional brake would be such a device, but its properties are strongly nonlinear and hard to control. Alternatives are therefore highly interesting. The team of COLGATE showed in [170] how to increase the impedance of an electronic - commutated electrodynamic actuator, whereby two windings were bypassed by a variable resistor. The mutual induction possible by this bypass damped the motor significantly. In [76] the team of HAYWARD went even further by implementing an eddy current break into a pantograph-kinematics (fig. 9.87). This break is a pure damping element with almost linear properties. By this method a controlled dynamic damping up to 250 Hz was achieved.

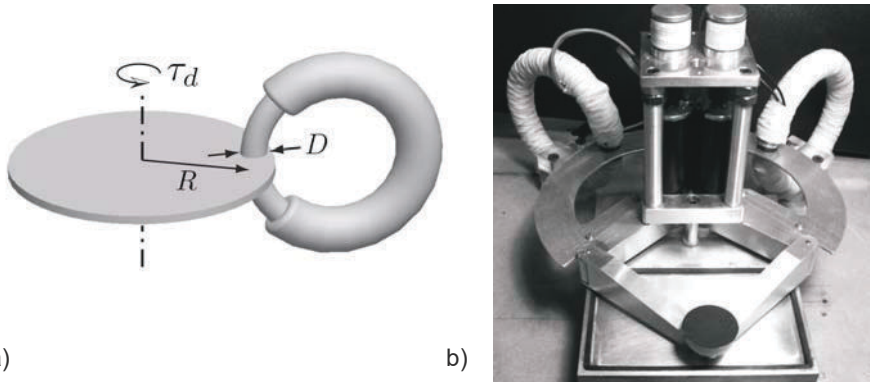


Fig. 9.87 Principle of eddy currents damping a rotating disc (a) and realization as a haptic device (b) by [76].

9.6.1.5 MagLev - Butterfly Haptics

In the 1990s the team of HOLLIS developed a haptic device [17] based on the electrodynamic actuation principle (fig. 9.88). Since recently, the device has been sold commercially by “Butterfly Haptics”. It is applied, e.g., to ongoing research projects on psychophysical analysis of texture perception. Six flat coils are mounted in a hemisphere with a magnetic circuit each. The combination of LORENTZ-forces of all coils allows an actuation of the hemisphere in three translational and three rotational directions. Via three optical sensors - each of them measuring one translation and one rotation - the total movement of the sphere is acquired. Besides the actuation within its space, the control additionally includes compensation of gravity with the aid of all six actuators. This function realizes a bearing of the hemisphere with Lorentz-forces, only. The air-gap of the coils allows a translation of 25 mm

and a rotation of $\pm 8°$ in each direction. As a consequence of the small mass of the hemisphere, the electrodynamic actuator principle as a drive and the abandonment of mechanical bearings, forces of a bandwidth of 1 kHz can be generated.

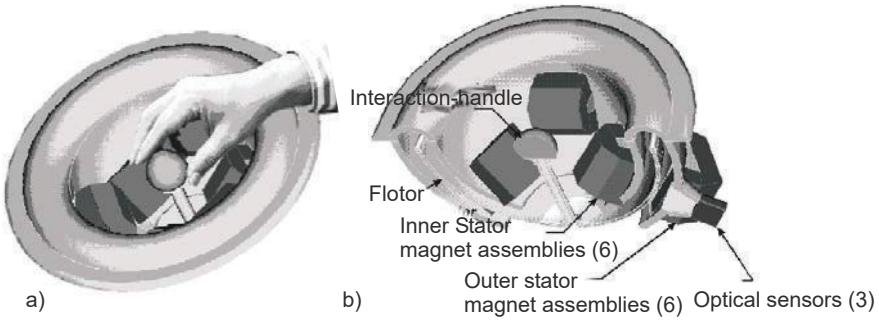


Fig. 9.88 The *MagLev* device with a hand resting on its handle (a) and the sketch of the internal structure (b) [17].

9.6.2 Haptic-Tactile Devices

Haptic-tactile devices of this category are intelligent combinations of well-known actuator principles of haptic systems with either high position resolutions or extraordinary, dynamic properties.

9.6.2.1 Thermo-Pneumatic

A classic problem of tactile pin-arrays is given by the high density of stimulator points to be achieved. The space below each pin for control and reconfiguration of the pin's position is notoriously finite. Consequently, a large number of different designs has been tested until today. In [272] a thermo-pneumatic system is introduced (fig. 9.89) based on tubes filled with a fluid (methyl-chloride) with a low boiling point. The system allows a reconfiguration of the pins within 2 seconds. however, it has high power requirements, although the individual elements are very cheap.

9.6.2.2 Aggregate State for Fixation

An alternative approach is pursued in [182] by changing the state of matter from fluid to fixed (fig. 9.90). The movement of the pins is brought about by air pressure.

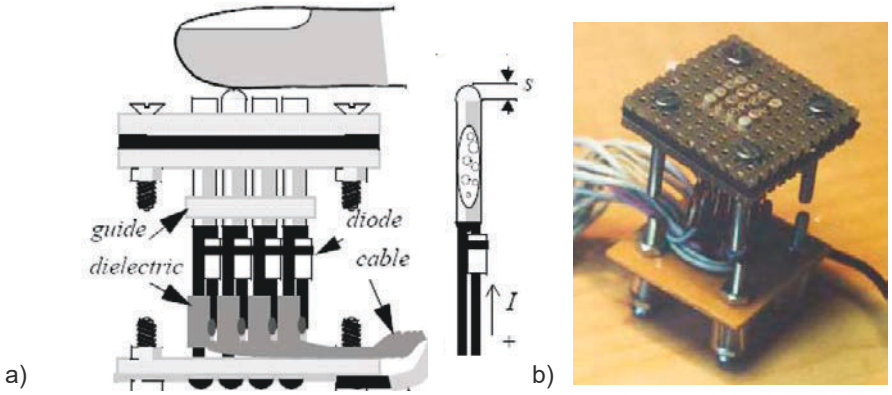


Fig. 9.89 Thermo-pneumatic actuation principle in a schematic sketch (a), and as actual realization (b) [272].

One part of each pin is sticking in its own pot filled with a metal (U-alloy) of a low melting point. Each pot can be heated separately. In the fluid state of the metal the pin is moved by air pressure, in the fixed state the pin is immovable. Based on this principle, array resolutions of 2 mm are quite easily achievable.

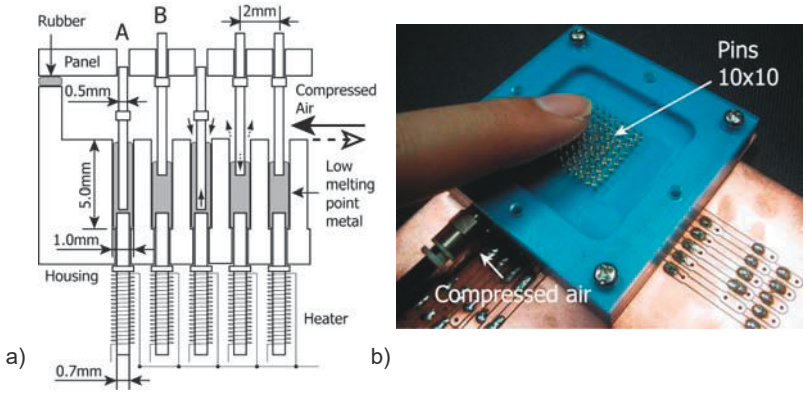


Fig. 9.90 Tactile display with a fixation of the individual pins by changing the aggregate state of a metal: cross-section (a), and realization (b) [182].

9.6.2.3 Fishbone-Illusion

Similar to the visual and acoustic perceptions, the haptic sense can be subject to impressive illusions. Among the best known illusions is the one consequently researched by NAKATANI [183]: generating the perception of concave or convex lines when touching completely different geometrical structures. This "Fishbone"-illusion (fig. 9.91) - according to today's knowledge - is based on the effect that shear forces are coupled to neighboring skin areas when touching these structured surfaces, which show a similarity to tensions and strains generated by real shapes. HAYWARD, an enthusiastic lateral thinker of haptic science, only recently published an overview of the best-known tactile illusions [89], suggesting a structure of the individual effects and their dependence of each other.

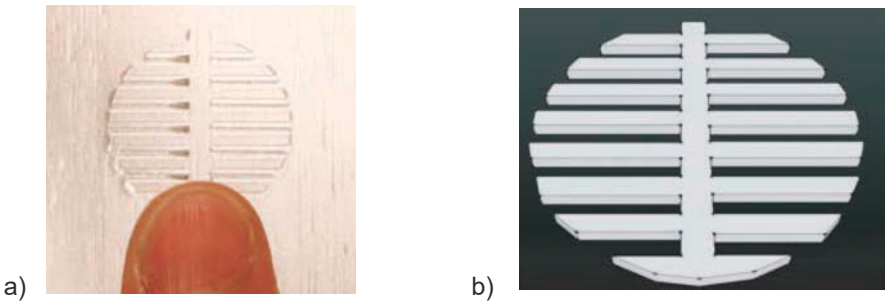


Fig. 9.91 Fishbone-structure generating the illusion of a line with concave bending: in real proportions (a), and as scheme (b) [183].

9.6.2.4 Piezoelectric Texture Actuator

Besides an application in Braille-related tasks, the design of tactile displays is relevant for texture perception, too. At the World Haptics Conference in 2007, WINFIELD impressively demonstrated a simple tactile texture display on the basis of a piezoelectric disk in resonance mode [288]. By variation of frequency and amplitude it was possible to put the finger touching the disk into different degrees of levitation, thus actually affecting the frictional coefficient between both. By the aid of an optical tracking right above the disc, and with a corresponding modulation of the control signal, perceivable textures with touchable, spatial resolutions were generated (fig. 9.92).

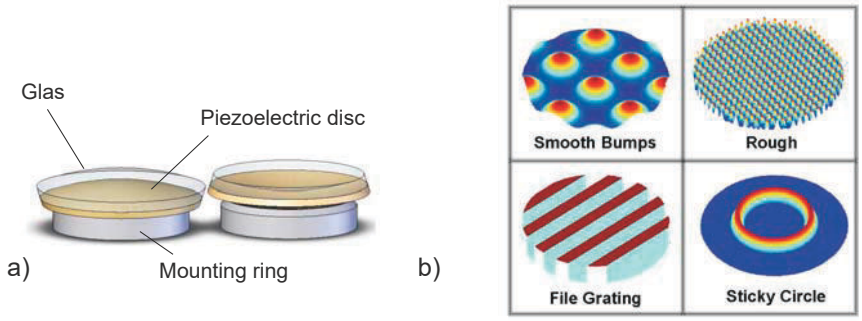


Fig. 9.92 Piezoelectric disc on a glass substrate (a), and variants of displayable textures (b) [288].

9.6.2.5 Volume-Ultrasonic Actuator

IWAMOTO built tactile displays (fig. 9.93) which are made of piezoelectric actuators and are actuated in the ultrasonic frequency range. They use sound pressure as a force transmitter. The underlying principle is given by generating a displacement of the skin and a corresponding haptic perception by focused sound pressure. Whereas in the first realization an ultrasonic array had been used to generate tactile dots in a fluid [112], later developments used the air for energy transmission [113]. The pressures generated by the designs (fig. 9.93) provide a weak tactile impression, only. But especially the air-based principle works without any mechanical contact and could therefore become relevant for completely new operation concepts combined with gesture recognition.

9.6.2.6 Perception of Surface Adhesion - Stickiness Perception

A seldom addressed effect of every day interaction is based on the perception of adhesion between skin and object. Adhesion may result from pollution of surfaces, but does frequently happen with soft materials like silicone, rubber, or, of course, glues. A device generating the effect of adhesion is presented in [291], and used with characteristic adhesion curves measured before. The device (fig. 9.94) is based on three vacuum pumps, able to generate negative pressure below the finger in three independent elliptic rings. By the variation of the vacuum within the concentrically located cylinders, it is possible to retain the finger in three discrete steps. The combined perception is very similar to the perception of sticking to surfaces with varying adhesiveness.

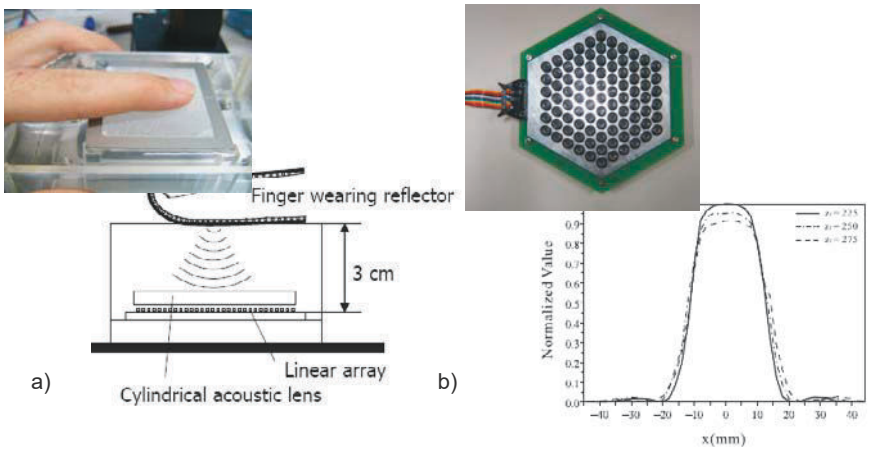


Fig. 9.93 Tactile display based on ultrasonic sound pressure transmitted by a fluid [112](a) or as array of senders for a transmission in the air [113] (b).

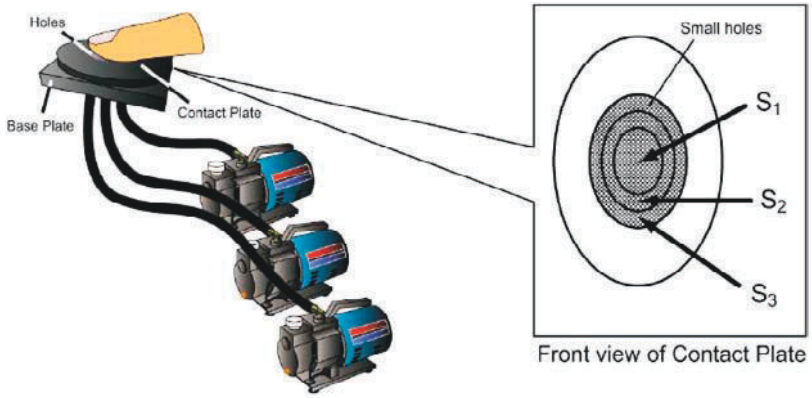


Fig. 9.94 Sketch of the device for generating a sensation of surface adhesion [291].

9.6.2.7 Tactile Contact Lens

As tactile perception mainly bases on the perception of tensions and strains coupled to the skin, it is possible to use levers between object and finger to magnify fine surface structures on the skin. This principle had been used in [136] to manufacture a tactile contact lens as shown in figure 9.95. The tactile contact lens is a purely mechanical device, using levers attached to a membrane in order to to extend the strain in the skin resulting from the touched structures.

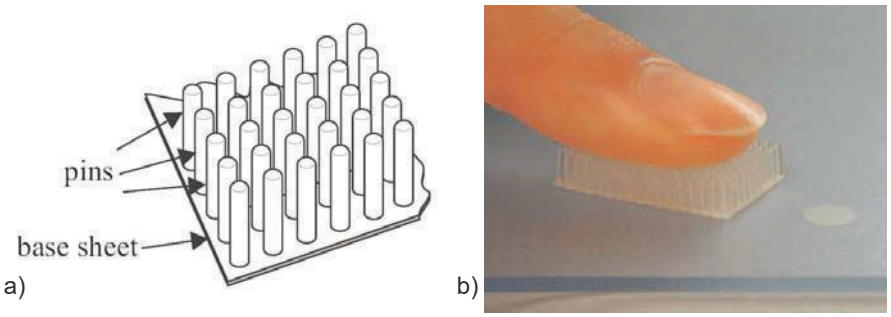


Fig. 9.95 Tactile contact lens for magnification by shear-tensions applied to the skin's surface: sketch (a), and realization (b) [136].

9.6.2.8 Tactile Markers by Tangential Forces

The findings that especially tangential forces and surface tensions are responsible for tactile perception, suggest further research whether controlled shear forces are able to generate perceptions comparable to a levitated pin on an otherwise flat surface. The team of HAWYARDhas been developing shear-force displays for Braille-related applications for several years now. The system shown in [161] (fig. 9.96) uses piezoelectric bending actuators for the generation of tangential forces. The piezoelectric bending actuators are additionally placed in a half-circuit to improve the coupling to the skin. The generated tactile impression is very realistic and can hardly be distinguished from a moving pin below the finger surface. A related tactile illusion is known as "hair-comb-effect".

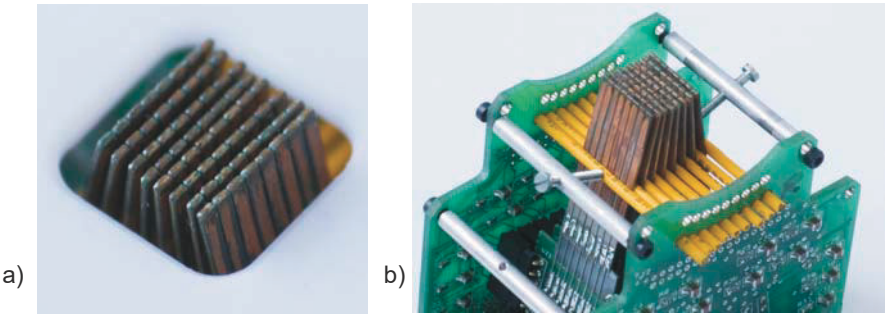


Fig. 9.96 Array made of piezoelectric bending actuators for shear-force generation in a planar (a), and in a curved "STReSS^2" design[161].

9.6.2.9 Haptic Display for Free-Space Motions

In case of the interaction with large virtual worlds it is frequently necessary to design devices which are worn on the body. An interesting solution has been shown in [256], generating a tactile sensation with belts (fig. 9.97) at the palm and at each finger. The underlying principle is based on two actuators for each belt, generating a shear force to the skin when being operated in the same direction, and a normal force when being operated in the opposite direction. This enables us to provide tactile effects when grasping or touching objects in a virtual world, but without the corresponding kinaesthetic effects. Nevertheless there are positive user responses to such devices.

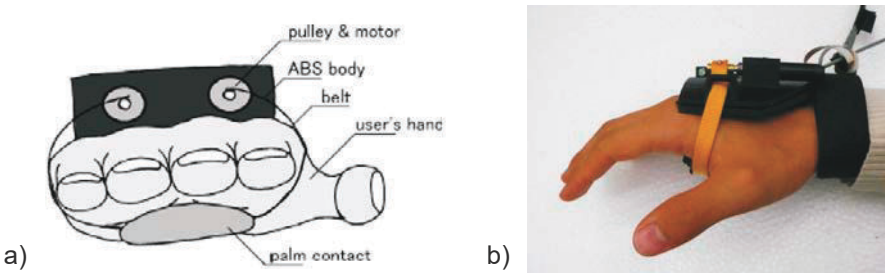


Fig. 9.97 Worn tactile device for the generation of normal and shear-forces: sketch for a palm-version (a), and the actual device (b)[256].

9.6.2.10 Electro-Tactile

As haptic receptors can be stimulated electrically, it is not far-fetched to design haptic devices able to provide low currents to the tactile sense organs. The design of such devices can be traced back to the 1970s. One realization is presented in [124] (fig. 9.98). Electro-tactile displays do work - no doubt - however they have the disadvantage also to stimulate noci-receptors for pain sensation beside the mechano-receptors. Additionally, the electrical conductivity between display and skin is subject to major variations. These variations are inter-person differences due to variations in skin-thickness, but they are also a time-dependent result of electro-chemical processes between sweat and electrodes. The achievable tactile patterns and the abilities to distinguish tactile patterns are subject to current research.

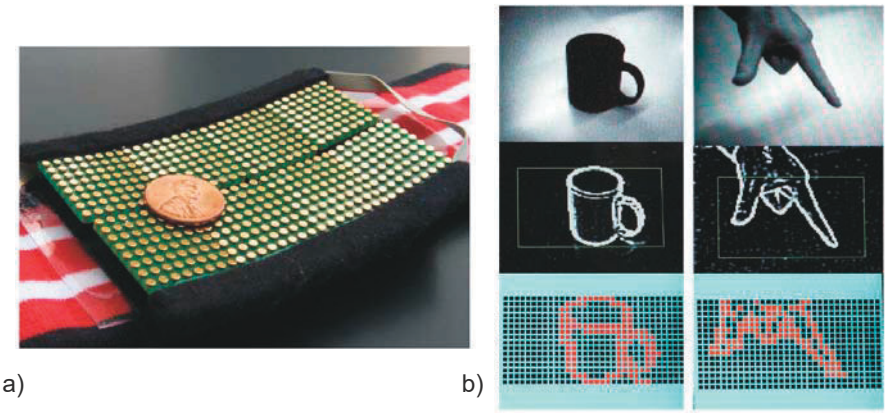


Fig. 9.98 Electro-tactile display worn on the forehead: Electrodes (a), and edge recognition and signal conditioning principle (b) [124].

Chapter 10
Force Sensor Design

JACQUELINE RAUSCH

The following chapter deals with selection and design of force sensors, which are implemented in haptic systems. In section 10.1 fundamental problems are discussed, which are the basis of every sensor design process. A selection of factors, which have to be taken in account, is made in section 10.1.5. After a short introduction in basic transfer properties, the sensor characteristics are analyzed according to haptic aspects and complemented by application examples.

10.1 Constraints

The topology of haptic systems significantly influences the design of force sensors. The application of the haptic device itself has an extraordinary relevance. All systems have in common that an user mechanically contacts objects. On this has to be clarified, which use of the device is intended, e.g. if it is going to be a telemanipulator for medical purposes, or a CAD tool with force feedback. The mechanical properties of the user itself and in case of telemanipulation systems the mechanical properties of manipulating objects, have to be analyzed for the sensor development. All these factors are going to be discussed within this section.

10.1.1 Topology of the Device

The application itself appoints the topology of the haptic device. Taking control engineering aspects into account haptic systems can be basically classified in four

T.A. Kern (ed.), *Engineering Haptic Devices*, DOI 10.1007/978-3-540-88248-0_10,
© Springer-Verlag Berlin Heidelberg 2009

types, which are discussed in chapter 5. In the following these topologies are analysed referring to the measured values:

- Open-loop control of impedance: Measurement of user movements (velocity or displacement), feedback of a force
- Closed-loop control of impedance: Measurement of both user movements and interaction, feedback of a force
- Open-loop control of admittance: Optional measurement of user force, feedback of a position
- Closed-loop control of admittance: Measurement of both user force and movements, feedback of a position

In case of an open-loop control only the mechanical properties of objects has to be taken into account for force sensor design, independent if objects are physical or virtual ones. In case of haptic simulators like flight simulators virtual objects are acting. The mechanical properties are often stored in look-up tables and force sensors are dispensable. In case of telemanipulation systems like e.g. da Vinci, the endeffector of the haptic system interacts with physical objects. Their mechanical properties have to be measured with capable force sensors. Most of the telemanipulation systems are impedance controlled. In case of closed-loop control the mechanical impedance of both user and manipulating object are considered. Designing closed-loop impedance controlled systems force sensors have to be integrated into the device detecting the user force. Designing closed-loop admittance controlled systems the user movements have to be measured using e.g. a velocity sensor (chapter 7).

Consequently, the measuring object can be both the user it self and a real, physical object. Beside its mechanical properties the modality of the interaction with the haptic systems has to be analyzed to identify fundamental sensor requirements like dynamic bandwidth, nominal load and resolution. The main factors influencing the sensor design are both the contact situation and the objects' mechanical properties. In the following, they are analyzed by examining separately mechanical properties and texture of the objects' surface.

10.1.2 Contact Situation

It is necessary to distinguish between the user of the haptic system and the physical object due to the different interaction modalities identifying mechanical properties. If the user is the "measuring object", interaction forces have to be measured. Universally valid conclusions relating to amplitude, direction and frequency of the acting force cannot be done. The mechanical impedance depends on the manner of grasping the device, age and gender of the user itself (section 4.2). In section 4.2.1 the manners of grasping are classified: power-grasps, precision-grasps and touch-grasps. In case of power- and precision-grasps finger or palm are used as counter

bearing, which results in high absolute value of force up to 100 N [34, 88] and a stiffer contact.

Additionally, the direction of the force vector has to be taken into account. Depending on application of the haptic device and manner of grasping up to six degrees of freedom result - three force components and sometimes three torques. Neglecting the torques between user and device three components of force have to be measured. If the user is in static contact with the device's handheld, measuring normal force components with respect to the orientation of the contact plane is sufficient. If the user is exerting relative movements to the device, also shear forces occur and three components have to be measured.

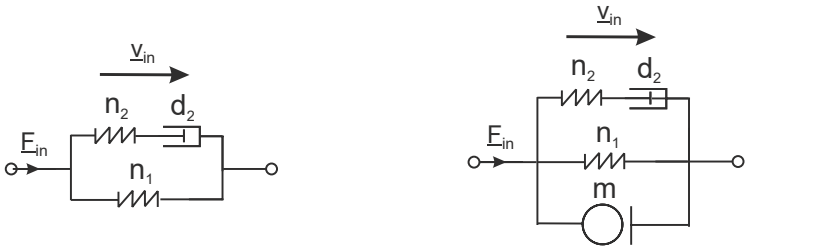
Considering the frequency dependency of humans' haptic perception, both static and dynamic signal components have to be considered equally (chapter 3). The lower cut-off frequency of haptic devices tends to quasi-static action at almost zero Hertz, which may happen when a device is held without movement in free space. If the force signal is subject to noise or even the slightest drift, the haptic impression will be disturbed soon (compare perception thresholds in section 3.1). Manner and pre-load of grasping affect the upper cut-off frequency of the sensor. In case of power- and precision-grasps, the absolute value of force achieves higher values which results in an upper cut-off frequency being $<<$ 10,000 Hz [1]. Values of about 300 Hz are sufficient (section 3.1). Within the contact grasps the pre-load is much lower than before enabling high frequency components to be transmitted directly to the skin up to a range of approximately 1,000 Hz. In case of telemanipulation systems the endeffector interacts with a real, physical object. The assumptions made for the measuring object "user" can partially be transferred to this situation. Following Newton's law *actio et reactio*, the absolute value of force depends on intensity and way of interaction. Possible examples are the compression and lift of objects with a gripper, or exploration with a stick. For telemanipulation systems in minimally invasive surgery the absolute values of force range from 1 N to 60 N (comp. e.g. [210]). The most promising approach is given by analyzing the intended application within preliminary tests and derivation of a model. The dynamics of the interaction, especially of the upper cut-off frequency, is dominated by the mechanical impedance of the object itself, which will be described within the following section.

10.1.3 Mechanical Properties of Measuring Objects

As stated for the user in chapter 5, the mechanical impedance of an object can be subdivided into three physical actions: the elastic compliance n, the damping d and the mass m. In case of rigid objects made of e.g. metal or ceramics, the property of elasticity is dominant. The interaction between haptic systems and objects can be considered as a rigid contact. Consequently, the force signal includes high-

[1] Please note: as stated in section 3.1, vibrotactile perception enables to feel frequencies of up to 10 kHz.

frequency components. The upper cut-off frequency should take a value of at minimum 1,000 Hz, to make sure to cover all dynamics responsible for haptic perception. Soft objects, such as silicone or viscera have a viscoelastic material performance. Following KELVIN viscoelastic behavior can be simulated by a network made of elastic compliances n_i and damping elements d_i, such as masses m_i. Using such an equivalent network dynamic effects like relaxation and creeping can be modeled (fig. 10.1 and fig. 10.2).



(a) Kelvin model modeling dynamically effects

(b) Kelvin model extented by objects weight for calculating the resonance frequency

Fig. 10.1 Kelvin model (standard linear solid) modeling viscoelastic behavior of objects. For calculating the resonance frequency a mass element has to be added. Adding further damping and spring elements the dynamic behavior of every object material can be modeled.

First of all the elasticity of mesuring objects has to be analysed during the sensor design. An arithmetic example in section 3.3.1 compares the different cut-off frequencies of materials. For soft materials such as rubber the upper cut-off frequency takes values below 10 Hz. During interaction with soft materials mainly low frequency components appear. The upper cut-off frequency is defined by the interaction frequency of 10 Hz at maximum [128, 204, 233]. If the measuring object is a soft one with embedded rigid objects, like e.g. tumors in soft body tissue, an upper cut-off frequency of about 1,000 Hz should be realized. To get more precise information about the frequency requirements it can hardly be done without an analysis of the interaction object. For a first rule of thumb calculated cut-off frequency as derived in section 3.3.1 are sufficient. In case of doubt, the frequency range of the sensor should always be oversized, not to already loose relevant haptic information already at this very first point in the processing chain.

As the dynamics does, the required force resolution is also depending on a physiological value. The *Just Notable Difference (JND)* lies in the range between 5 to 10 % of the absolute force value (chapter 3.1). From the JND the sensor characteristics of measurement uncertainty can be derived. If realized as a constant value - which is common to many technical sensor solutions - 5 % of the lowest apearing value of force should be chosen to prevent distortion of the haptic impression of the object. Nevertheless, there is no actual requirement for haptic applications to

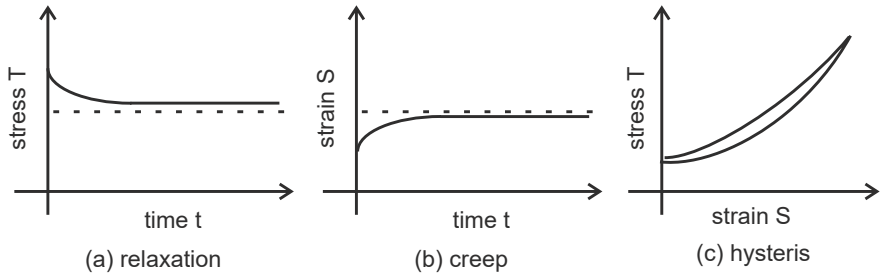


Fig. 10.2 Visualization of viscoelastic phenomena relaxation, creep and hysteresis.

have a constant or even linear sensor resolution. With telemanipulation systems the interaction of the haptic system and real, physical objects is the main application. Depending on the type of interaction, frequently the surface structure of objects, the so called texture become equally or even more important than the object's elastic compliance. The resulting challenges for sensor development are discussed within the following subsection.

10.1.4 Texture of Measuring Objects

The properties, which are relevant for the human perception of texture, are the geometrical surface structure on the one hand (e.g. the wood grain), on the other hand some kind of "frequency image" generated by the geometrical structure in the (vibro-)tactile receptors when being touched by skin. To detect the surface structure of an object, the variation of force against the contact area can be derived. For **static measurement** sensor arrays of single-component force or pressure sensors are a common technical solution. These arrays are put on the object. The structure of the object generates different values of contact forces, providing a force distribution on the sensor surface. The size of array and individual array elements can not be defined in general, but is depending on the smallest detectable structure on the measurement object itself. In case of the static measurement sketched above, number and size of the sensor array elements should be dimensioned slightly smaller than the minimum structure of the measuring object. The size of each element may be not larger than half of the size of the smallest structure to be measured. However, even fulfilling this requirement aberration will appear. Figure 10.3 shows that in case of the width of the sensor element being larger or identical to the smallest structure the distance between the elements is detected smaller than in reality. With n sensor elements the width of the structure element is replayed to $\frac{n+1}{n}$ and the distance to $\frac{n-1}{n}$. If the number of sensor elements per surface area increase, the aberration is diminishing and

the structure is approximated more realistic (fig 10.4). However, with the number of elements the effort of signal conditioning and analysis is increasing.

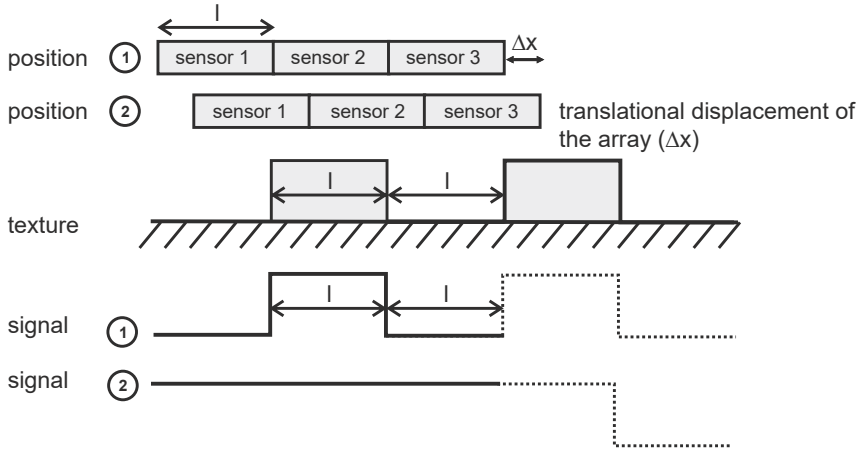


Fig. 10.3 Illustration of static and spatially resolved force measurement using as 3x*n* array. One sensing element has the same dimension like a texture element. At position 1 the array is optimally placed. If the array is shifted about Δx to position 2 the texture is incorrectly detected.

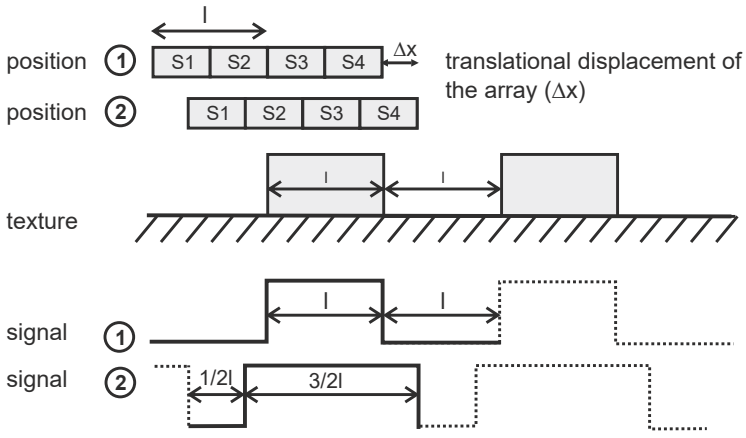


Fig. 10.4 Illustration of static and spatially resolved force measurement using as 6x*n* array. The size of one sensing element is half of a texture element. At position 1 the array is optimally placed. In case of any other position an aberration occurs. The aberration decreases with increasing number of sensing elements in an appropriate array.

Beside the described aberration an additional disadvantage of static measurements is given by the fact, that the knowledge of the texture is not sufficient to get information about the object's material. The complete haptic impression needs frequency information depending on the elastic properties of texture and surface friction too. To gain these data, a relative movement between object and haptic system should be performed, to measure the texture **dynamically and spatially**. Depending on velocity of the relative movement and speed of the signal detection algorithms the spatial resolution can be multiplied using the same number of sensor elements as in the example shown before. Even the use of sensor array with a simultaneous detection of multiple points becomes unnecessary. With knowledge about the exploration velocity and its direction the information can be put into relation to each other. For texture analysis multi-component force sensors should be used, as especially the combined forces in the direction of movement and normal to the surface contribute to haptic perception [184]. This dynamic measurement principle is comparable with the intuitive exploration made by humans: To gain the texture of an object humans gently touch and stroke over its surface. The surface structure excites the fingerprint to oscillate and the vibrotactile sensors acquire the frequency image. The absolute value of normal forces reached during such explorations are in a range of 0.3 to 4.5 N [35]. The force resolution is defined - as stated earlier - by the *JND*. The haptic information about the texture is included into the high-frequency components of the signal. For haptic applications the maximum frequency should be located at 1,000 Hz. The absolute value of nominal force should be chosen depending on the elastic compliance of the object. If the object is softer, the absolute value can be chosen lower as surface structures will deform and cannot be detected anymore. To be able to measure equally good at soft and rigid objects, the nominal force should take values $\leq 4.5\,\mathrm{N}$. CALDWELL [35] for examples decided to use $F = 0.3\,\mathrm{N}$.

10.1.5 Selection of Design Criteria

Following the description of the most relevant constrains, the limiting factors for a sensor design in haptic applications can be found in physiological values. Nominal force, force resolution, covered frequency range and measurement uncertainty can be derived from humans' haptic perception. For a quantitative analysis of these requirements the contact between measuring object and force sensor is to bring into focus. Measurement range and number of detectable force components are defined by the application and the structure of the device. The geometrical dimensions and other mechanical requirements are given depending on the point of integration into the haptic system. The diagram displayed in figure 10.5 visualizes the procedure how to identify the most important requirements for the sensor design

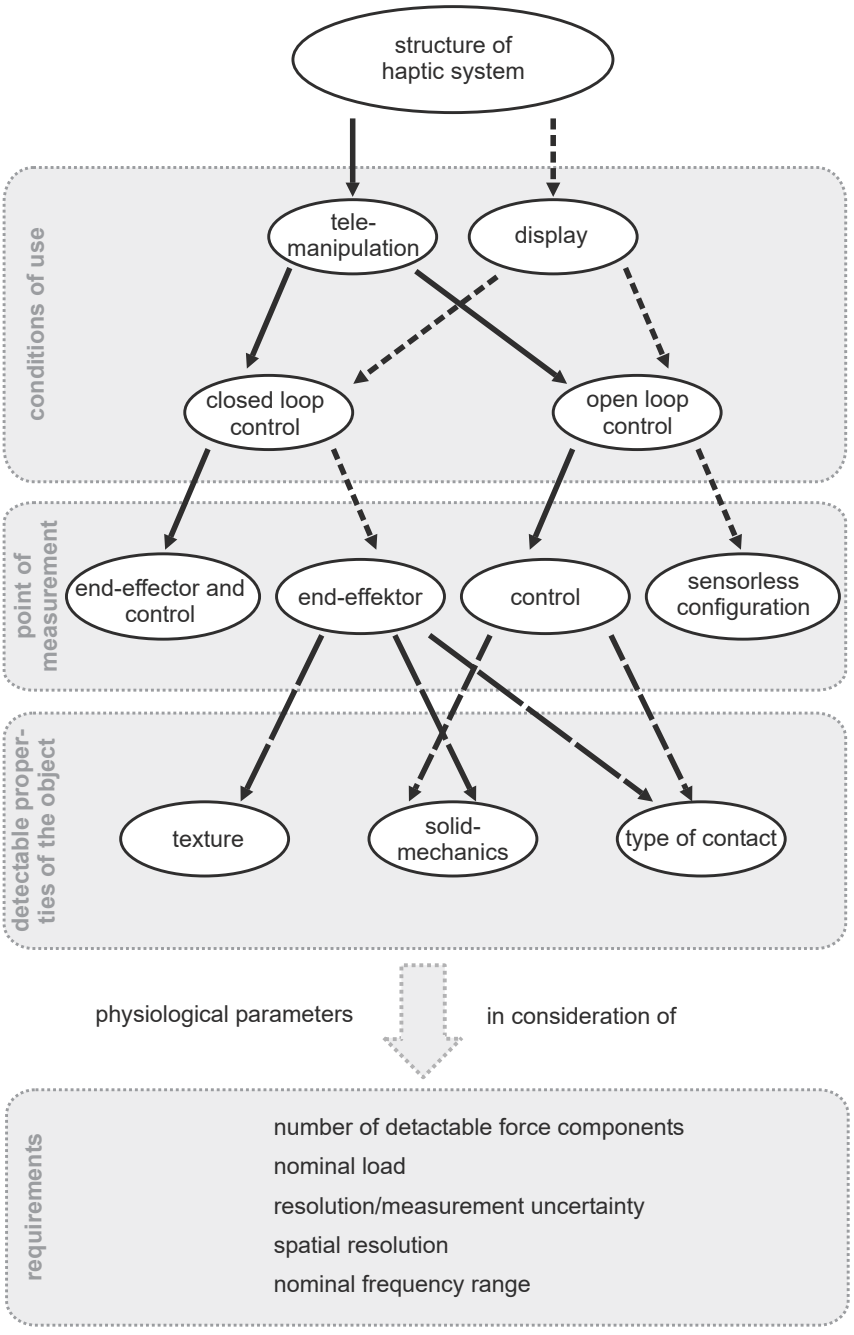


Fig. 10.5 Tree diagram to identify the principle requirements on haptic force sensors. Beside mechanical characteristics of the object also physiological parameters of human haptic perception have to be considered.

10.2 Sensing Principles

Within the previous section the most important criteria for the design and development of a haptic sensor were named and introduced. Section 10.3 will summarize the major requirements once again in tabular form. To help with the choice of a suitable sensor principle, the variants according to figure 10.6 are presented in this section. Beside the established measurement elements, such as resistive, capacitive, optic or piezoelectric ones, other "exotic" sensor designs based on electro-luminescence or active moving coils are discussed too.

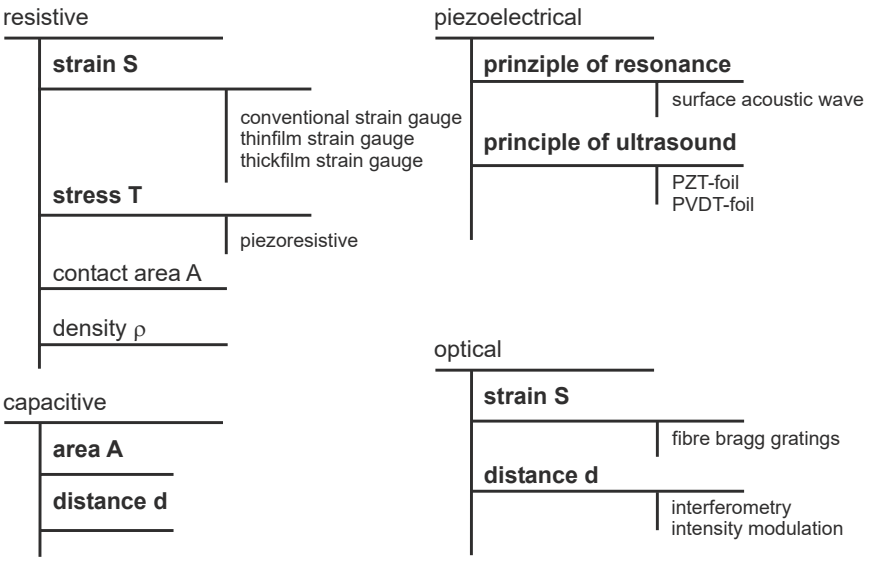


Fig. 10.6 Overview of established measurement principles for detecting forces in haptic systems. Furthermore active sensor systems are also discussed.

Most sensor principles are active transformers using the displacement principle for force measurement, which means that elasto-mechanic values such as stress or strain are detected and the corresponding force is calculated. Sensors which belong to the group of active transducers are resistive, capacitive, optic and magnetic ones working according to the displacement-principle too. Piezoelectric, electro-dynamic or electrostatic sensors belong to the group of passive transducers. After a short introduction in elasto-mechanics each sensing principle will be discussed according to its operating mode and several applications will be presented. All sensor principles will be estimated according to their applicability for kinaesthetic and tactile force measurement, and put into relation to the requirements known from chapter 6. At the end of this chapter a ranking method for the selection of suitable sensor principles will be given.

10.2.1 Basics of Elasto-mechanics

As aforementioned, a reasonable number of sensor principles base upon elasto-mechanics. This section will summarize the fundamental knowledge which is necessary for sensor design. A force is exerted to an elastic body. Depending on the value of force the body deforms elastically. Internal stress T occurs resulting in a shape change of the body - the strain S. Stress and strain are correlated by specific material parameters, the so called elastic moduli s_{ij}. For a better comprehension a short gedankenexperiment will be performed [220]. If a volume element ΔV is cut from an object under load (fig. 10.7), the substitutive forces ΔF will act upon the surfaces of the cuboid to keep the state of deformation. Due to the required state of equilibrium the sum of all forces and torques acting upon ΔV must equal zero.

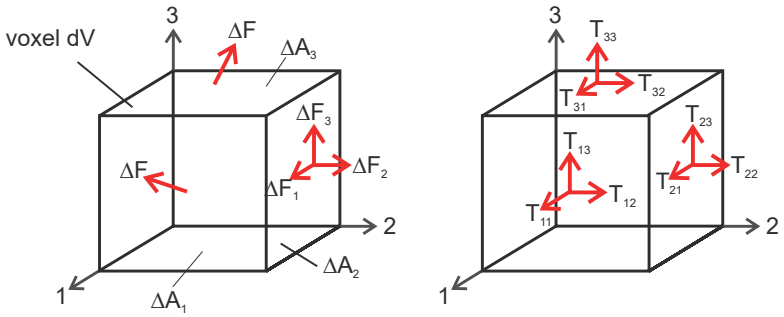


Fig. 10.7 Voxel dV of an elastic object. Due to deformation internal stress occurs which an be described by element T_{ij} of the stress tensor [220].

Subdividing the force ΔF in its three components ΔF_1, ΔF_2 and ΔF_3, just those components remain orthogonal to the surface elements ΔA_j. The quotient of the acting force component ΔF_i and the corresponding surface element ΔA_j results in a mechanical stress T_{ij}. Following the balance condition $T_{ij} = T_{ji}$ six independent tension components remain, resulting in the stress tensor. The tensor elements can be factorized into normal (stress parallel to the surface normal) and shear stress components (stress orthogonal to the surface normal). Analyzing the volume element ΔV before and after load, a displacement of the element ΔV with relation to the coordinate system $\langle 123 \rangle$ such as a deformation happens. The sides of the cube change their lengths and are not orthogonal to each other anymore (fig. 10.8).

For the description of this shape change a strain S_{ij} is introduced. This strain is a tensor too, consisting of nine elements (equ. 10.1)

$$\begin{pmatrix} d\xi_1 \\ d\xi_2 \\ d\xi_3 \end{pmatrix} = \begin{pmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{pmatrix} \cdot \begin{pmatrix} \Delta x_1 \\ \Delta x_2 \\ \Delta x_3 \end{pmatrix} \tag{10.1}$$

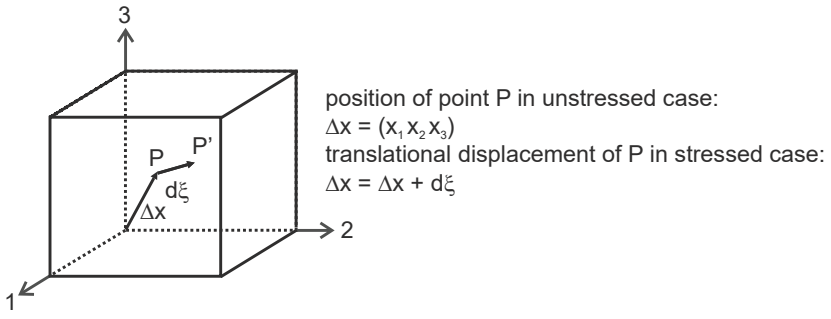


Fig. 10.8 The displacement of point P to P' due to application of force visualizes the state of strain [220].

Due to the volume constancy the following correlation can be defined

$$S_{ij} = S_{ji} = \frac{1}{2} \cdot \left(\frac{\delta \xi_i}{\delta x_j} + \frac{\delta \xi_j}{\delta x_i} \right) \tag{10.2}$$

and the matrix can be reduced to six linear independent elements. The normal strain components act parallel to the corresponding normal to the surface. This is resulting in a volume change. The shear components, acting normal to the surface, describe the change of the angle between the borders of the volume element.

In case of isotropic materials, such as e.g. metals or Al_2O_3 ceramics, the interconnection between the shape change mentioned before and the mechanical strains can be formulated as follows:

$$\begin{pmatrix} S_1 \\ S_2 \\ S_3 \\ S_4 \\ S_5 \\ S_6 \end{pmatrix} = \begin{pmatrix} s_{11} & s_{12} & s_{12} & 0 & 0 & 0 \\ s_{12} & s_{11} & s_{12} & 0 & 0 & 0 \\ s_{12} & s_{12} & s_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & 2(s_{11}-s_{12}) & 0 & 0 \\ 0 & 0 & 0 & 0 & 2(s_{11}-s_{12}) & 0 \\ 0 & 0 & 0 & 0 & 0 & 2(s_{11}-s_{12}) \end{pmatrix} \cdot \begin{pmatrix} T_1 \\ T_2 \\ T_3 \\ T_4 \\ T_5 \\ T_6 \end{pmatrix} \tag{10.3}$$

The six independent strain resp. stress components are summarized as a vector for simplification. The components with the index 1, 2 and 3 mark the normal components, those with indices 4, 5 and 6 the shear components [284]. The parameter s_{ij} are regardless of direction. From them the known elasticity modulus E and shear modulus G can be derived:

$$s_{11} = \frac{1}{E}, s_{12} = \frac{\nu}{E}, \frac{1}{G} = 2(s_{11} - s_{12}) = \frac{2}{E}(1+2\nu) \tag{10.4}$$

ν marks the so called POISSON ratio, which is material dependent. With metals ν achieves values between 0.25 and 0.35. For homogeneous materials equation 10.3

can be reduced to a linear correlation $T = E \cdot S$ as a good approximation. For anisotropic materials such as silicon or quartz the elastomechanic properties are dependent on the orientation of the coordinate system (comp. section 10.2.3), resulting in a matrix of elastic coefficients with up to 21 elements. For further reading on elasto-mechanics e.g. [220] can be recommended.

10.2.1.1 Example "Beam Bending"

If a force vector is exerted to the tip of a beam bender made of isotropic materials and clamped on one side (fig. 10.9), a bending moment M_B acts upon the beam. The

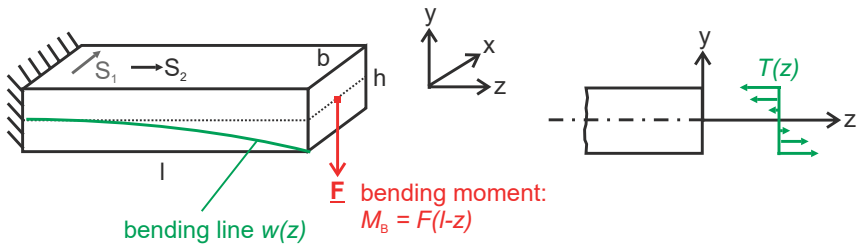


Fig. 10.9 Behavior of a bending beam, the right-hand detail shows the stress distribution along the profile.

mechanical stress components $T(y)$ are linear distributed on the cross section and take values of $T(y) = c \cdot y$, whereas c is a proportional factor. The bending moment equals the integral of the stress $T(y)$ distributed on the cross section.

$$M_B = \int_A y \cdot T(y) dA = c \cdot \int_A y^2 dA \tag{10.5}$$

As the integral $\int_A y^2 dA$ equals the axial moment of inertia I, c is calculated as

$$c = \frac{M_B}{I}. \tag{10.6}$$

The resulting strain components S_1 and S_2 act transversally to the beam's surface. Within the elastic region strain component S_1 and the stress component T_2 are correlated via the elasticity modulus E

$$S_2 = \frac{T_2}{E} = \frac{M_B}{I \cdot E} = \frac{F \cdot (l - z)}{I \cdot E} \tag{10.7}$$

and therefore depending on the geometry of the cross section A of the beam, the position z at the beam's surface and the acting force F. For calculations of the strain component S_1 the transversal contraction has to be considered as follows

$$S_1 = -\nu \cdot S_2. \tag{10.8}$$

Further readings of elasto-mechanics, for example the calculations of deformation of fiber-reinforced composites, the works of SCHNELL [220], WERTHSCHÜTZKY [284] and BALLAS [8] are recommended.

10.2.1.2 Detection of Force

According to figure 10.9 acting forces can be measured evaluating both a resulting strain distribution on the surface and a displacement of the beam. Following the example above, the detection of strain S_2 can be derived using BERNOULLIs theory. On this strain components acting transversal to the surface can be neglected for slender and long beam-geometries. The stress or strain sensitive elements should be placed in such a way, that a maximum surface strain change can be detected. The correlations described above are examples for a clamped cantilever. Being able to measure more than one force component, a suitable deformation element has to be designed considering the elasto-mechanic correlations. Works of BRAY [28] can help designing such an element. The primary objective is to generate a strain distribution in loading case, which enables to deduce the force components. Typically, the interconnection between the force F_i and the electric signal v_i of the sensor element is given by a linear system of equations (e.g. [274]). Equation 10.9 shows an example for a three-axial sensor:

$$\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \cdot \begin{pmatrix} F_1 \\ F_2 \\ F_3 \end{pmatrix} \tag{10.9}$$

It can be assumed that all force components contribute to each individual voltage signal v_i. The elements a_i of the matrix can be found by calibrating the sensor. During the calibration process only one independent force component for each direction is applied to the sensor and the resulting voltage components are measured. After inverting the matrix $\mathbf{A}$ to $\mathbf{A}^{-1}$ the force vector can be calculated easily.

10.2.2 Resistive Strain Measurement

One of the most commonly used sensing principles for force sensing is based on the resistive detection of strain components occurring in a (measuring) object. With the resistive stress measurement a resistor pattern is applied on the bending elements surface. The resistors have to be located in the areas of maximum strain. For your information: The electrical resistance is defined via

$$R_0 = \rho \cdot \frac{l}{A} = \rho \cdot \frac{l}{b \cdot h} \tag{10.10}$$

ρ marks the specific resistance, l, b, h (length, width, height) define the volume of the resistor itself. The total differential shown in equation 10.11 gives the relative resistivity change resulting from the deformation.

$$\frac{dR}{R_0} = \underbrace{\frac{dl}{l} - \frac{db}{b} - \frac{dh}{h}}_{rel.\,volume\,changing} + \underbrace{\frac{d\rho}{\rho}}_{piezoresistive\,part} \tag{10.11}$$

The deformation causes on the one hand the change of the geometrical part $\frac{l}{A}$, and on the other hand the chance of the specific resistivity ρ. The material specific changes will be discussed within section 10.2.3. Using equation 10.12 the correlation between strain and relative resistivity change is formulated:

$$\frac{dR}{R_0} = \underbrace{\left(2 - \frac{d\,(N \cdot \mu)}{S \cdot N \cdot \mu}\right)}_{:=k,\,k-factor} \cdot S \tag{10.12}$$

whereas μ represents the electron mobility and N the number density of molecules. The change of the resistivity can be measured using a so called WHEATSTONE bridge circuit. This circuit is built of one to four active resistors connected in a bridge circuit and fed by a constant voltage or constant current (fig. 10.10). Equation 10.13 calculates the bridge 10.10(c) with the assumption, that the basic resistances R_{0i} equal the resistance R_0. The values of R_0 such as the k-factors are specific to the material and listed in table 10.1 (further informations e.g. [215]).

$$\Delta u = \frac{V_{cc}}{R_0 \cdot I_0} = \frac{1}{4} \cdot \left\{ \frac{r_1}{R_{01}} - \frac{r_2}{R_{02}} + \frac{r_3}{R_{03}} - \frac{r_4}{R_{04}} \right\} \tag{10.13}$$

The supply with a constant current has the advantage of the compensation of a temperature dependent drift of the measurement signal. More advanced information can be found in [44]. In case of metals a k-factor of approximately two occurs. The material specific component of metals is less important and affects the first decimal place only. In case of semiconductors and ceramic materials the material specific component is dominant. In case of semiconductor-strain gauges k-factor takes values up to 150. Using resistor pastes, applied in thick film technology on substrates[2], and poly-silicon layers, sputtered in thin film technology the material specific component is dominant. On this, k-factors achieve values of up to 18 in case of thick-film resistors and up to approximately 30 for thin-film resistors. Table 10.1 lists the k-factor for several materials usually used in strain measurement.

[2] For substrate material mainly (layer-) ceramics are used. Less frequent is the use of metals, as isolating layers have to be provided then

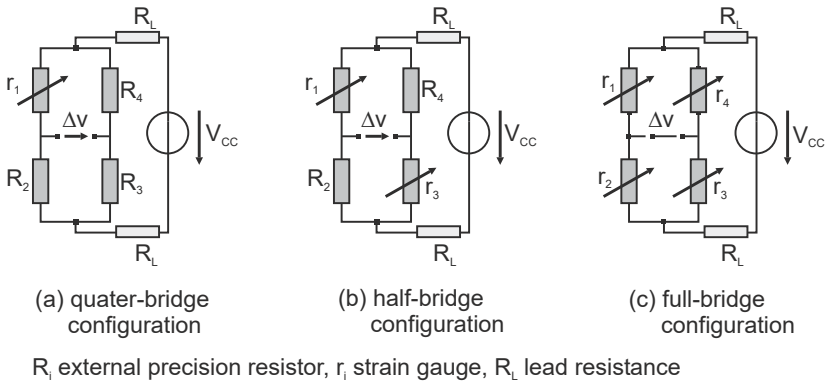


Fig. 10.10 WHEATSTONE bridge configurations for evaluating one up to four resistors.

Table 10.1 k-factors of important resistive materials [284]

technology	material	k-factor	basic resistance R_0 in Ω
foil strain gauge	*CuNi*	2.1	120, 350, 700
thickfilm	$Bi_2Ru_2O_7$	18,3	≈10k
thinfilm	*TiON*	4 ...5	≈5k
	Poly – Si	20 ...30	≈5k
silicon micromachining	*p – Si*	80.4	3k ... 5k

As mentioned earlier, strain gauges are manufactured in different technologies. The most commonly used variants are foil-strain gauges; thick- and thin-film manufactured measurement elements are found mainly in OEM-sensors and for specific solutions in automation industry due to the necessary periphery and the manufacturing process. Relevant literature can be found in the publications of PARTSCH [195] and CRANNY [42].

Foil strain gauges are multilayer systems made of metallic measurement grids and organic substrates. It is applied (fig. 10.11) and fixated on bending elements via cold hardening cyanoacrylate adhesive (strain analysis) or via hot hardening adhesives such as epoxy resin (transducer manufacture). These gauges are long-term stable, robust, and especially used for high-precision tasks in wind-tunnel-scales and balance sensors. Achievable dynamics, resolution and measurement range are solely depending on the deformation element. The minimum size of the individual strain gauges taken of the shelf is in the area of 3 mm width and 6 mm length. The measurement pattern itself is smaller in its dimensions. On this, it is possible to shorten the organic substrate to finally achieve 1.5 mm width and 5 mm length as a typical minimum size. If foil strain gauges are considered, the surface strains

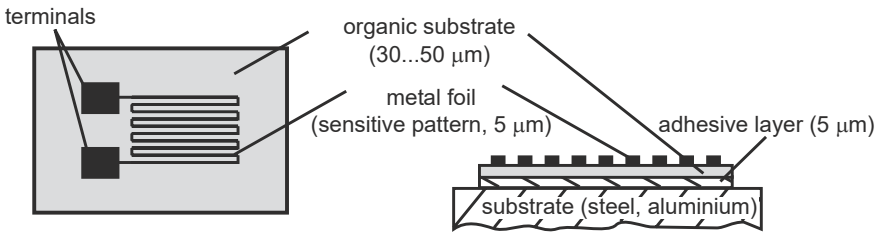


Fig. 10.11 Assembly of conventional strain gauges: the measuring grid is usually made of a textured metal foil. In case of special applications metal wires are applied.

resulting from the nominal load should be 1,000 μm/m for an optimum usage of the strain gauge. Many measurement pattern are applied for force and torque sensors. Figure 10.12 shows a selection of commercialized measuring grids ready for application on deformation elements.

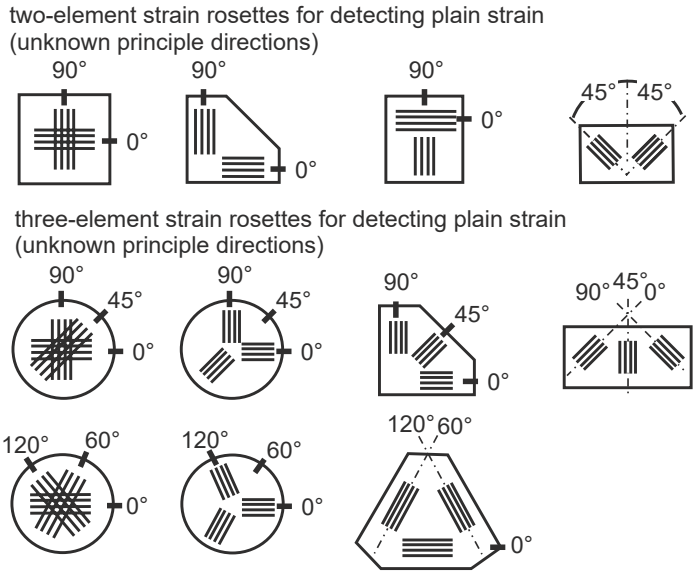


Fig. 10.12 Compilation of possible grid configurations of strain gauges. See also [284].

Beside resistive foil strain gauges, semiconductor strain gauges can be purchased. Their general design is comparable to conventional strain gauges, as the semicon-

ducting elements are assembled with organic substrates[3]. The measurement elements are used identical to foil strain gauges and are available in different geometrical configurations such as T-rosettes. Using measuring elements with a higher k-factor (table 10.1) the deformation elements can be designed stiffer, allowing smaller nominal strains. Such elements are especially relevant for the design of miniaturized sensors for haptic systems, as small dimensions and high cut-off frequencies have to be achieved. A commercially available example is the OEM-sensor *nano 17* from ATI (fig. 10.13). The strain elements are piezoresistive ones and their k-factor takes values of approximately 150. Due to the high potential for miniaturization and the manifold application in haptic systems, piezoresistive sensors - especially silicon sensors - will be discussed in an independent subsection.

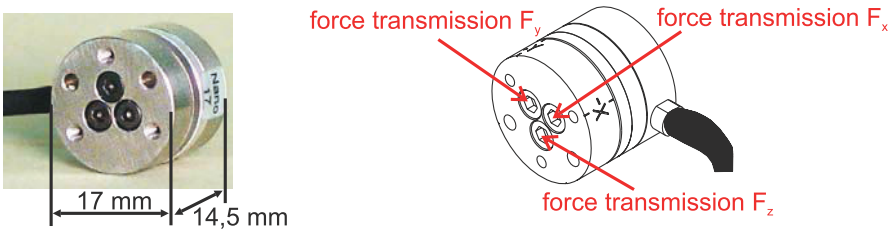


Fig. 10.13 Miniaturized force/torque sensor [6]. The resonance frequency of the sensor takes a value of about 7,2 kHz.

10.2.3 Piezoresistive Silicon Sensoren

Published by Charles S. SMITH in 1954 for the first time [237], semiconducting materials with a symmetric crystal structure such as silicon or germanium possess a change in their conductivity σ due to an applied force or pressure. In the following paragraphs this effect is discussed more deeply for monocrystalline silicon.

10.2.3.1 The Piezoresistive Effect

If a semiconducting material is deformed due to a load, stress components T_i are generated inside the material. For your information: Due to the anisotropic properties of the material the elasto-mechanic properties are depending on the position of the coordinate system, and consequently on the orientation of the crystal lattice.

[3] Also single semiconducting elements without organic substrate are available. They are highly miniaturized (width of about 230 μm, length of about 400 μm), but has to be insulated from the deformation element.

These stress components affect the electron mobility μ and - as a consequence - the specific resistivity ρ. ρ is a material specific value, characterized via the parameters electron mobility μ and number of charge carriers N (comp. section 10.2.1). By considering these parameters the correlation between relative resistivity change and the resulting strain tensor can be expressed to:

$$\frac{d\rho}{\rho} = \frac{dV}{V} - \frac{d(N \cdot \mu)}{N \cdot \mu}, \text{ with } \rho = \frac{V}{N \cdot \mu \cdot |q|}, \tag{10.14}$$

whereas V is the volume of the resistive area and $|q|$ is the charge of the particles.

Following the OHM's law the specific resistance ρ is connected by the vector $\mathbf{E} = (E_1; E_2; E_3)^T$ of the electrical field and the current density $\mathbf{J} = (J_1; J_2; J_3)^T$:

$$\begin{pmatrix} E_1 \\ E_2 \\ E_3 \end{pmatrix} = \begin{pmatrix} \rho_{11} & \rho_{12} & \rho_{13} \\ \rho_{21} & \rho_{22} & \rho_{23} \\ \rho_{31} & \rho_{32} & \rho_{33} \end{pmatrix} \cdot \begin{pmatrix} J_1 \\ J_2 \\ J_3 \end{pmatrix} = \begin{pmatrix} \rho_1 & \rho_6 & \rho_5 \\ \rho_6 & \rho_2 & \rho_4 \\ \rho_5 & \rho_4 & \rho_3 \end{pmatrix} \cdot \begin{pmatrix} J_1 \\ J_2 \\ J_3 \end{pmatrix} \tag{10.15}$$

Due to the symmetric crystalline structure of silicon[4] six independent resistive components ρ_i result, which are symmetrical to the diagonal of tensor ρ. With the matrix of piezoresistive coefficients π the influence of the six acting stress components T_i can be formulated. The cubic symmetry results in a reduction of the number of piezoresistive and direction dependent coefficients to three. By doping silicone with impurity atoms such as boron or phosphor areas of higher resistivity are generated. By influencing the type and the concentration of dopant the three π-coefficients can be influenced. Further information on doping can be found e.g. in [9].

$$\begin{pmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \\ \rho_4 \\ \rho_5 \\ \rho_6 \end{pmatrix} = \begin{pmatrix} \rho_0 \\ \rho_0 \\ \rho_0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} \pi_{11} & \pi_{12} & \pi_{12} & 0 & 0 & 0 \\ \pi_{12} & \pi_{11} & \pi_{12} & 0 & 0 & 0 \\ \pi_{12} & \pi_{12} & \pi_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & \pi_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & \pi_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & \pi_{44} \end{pmatrix} \cdot \begin{pmatrix} T_1 \\ T_2 \\ T_3 \\ T_4 \\ T_5 \\ T_6 \end{pmatrix} \cdot \rho_0 \tag{10.16}$$

For homogenous silicone with a small concentration of dopant the values in table 10.2 can be used.

Table 10.2 Piezoresistive coefficients of homogeneously doped silicon [9].

doping	N in $\frac{1}{cm^{-3}}$	ρ in Ωcm	π_{11} in $\frac{mm^2}{N}$	π_{12} in $\frac{mm^2}{N}$	π_{44} in $\frac{mm^2}{N}$
n-Si	$6 \cdot 10^{14}$	11.7	$-102.2 \cdot 10^{-5}$	$+53.4 \cdot 10^{-5}$	$-13.6 \cdot 10^{-5}$
p-Si	$1.8 \cdot 10^{14}$	7.8	$+6.6 \cdot 10^{-5}$	$-1.1 \cdot 10^{-5}$	$+138.1 \cdot 10^{-5}$

[4] face centered cubic

Depending on the angle between current density vector **J** and stress component T_i three effects can be distinguished. With the so called longitudinal effect the current i is guided parallel to the normal component of stress, with the transversal effect i is guided normal to the normal component of stress, and with the shear effect i is guided parallel or normal to the shear component of stress. Figure 10.14 visualizes the described correlations.

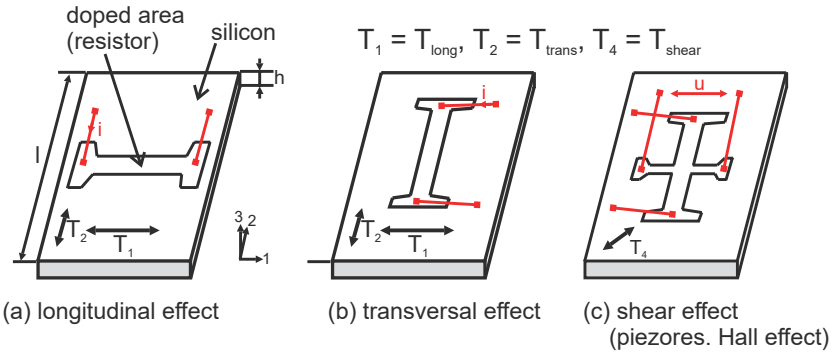


Fig. 10.14 Visualization of the the piezoresistive effects: longitudinal, transversal and shear effect in silicon [284]. Transversal and longitudinal effect is normally used for commercial silicon sensors.

For the resistivity change, depending on the orientation of the resistive area from figure 10.14, the following equation becomes valid:

$$\frac{d\rho}{\rho_0} = \pi_l \cdot T_l + \pi_q \cdot T_q \tag{10.17}$$

As a consequence longitudinal and transversal stress components are influencing the calculation of the resistivity change. Depending on the crystallographic orientation of the resistive areas the π-coefficient is built from the longitudinal- and transversal coefficient (table 10.2.3.1).

More advanced information for the design of piezoresistive silicon-sensors can be found in the publications of BAO [9], WERTHSCHÜTZKY [284] and STAVROULIS [240].

10.2.3.2 Examples of Piezoresistive Silicon Sensors

Piezoresistive silicon sensors for physical variables like pressure and force are commomly integrated in silicon deformation elements. For pressure transducers this manufacture is state of the art and for all pressure ranges sensor elements can be purchased. For example the company Silicon Microstructures Inc. (SMI) sells chips

Table 10.3 Compilation of π_l- und π_q-coefficients for selected resistor assemblies dependent on the crystallographic orientation [240].

surface orientation	longitudinal	π_l	transversal	π_q
(100)	[100]	π_{11}	[010]	π_{12}
	[110]	$\frac{\pi_{11}+\pi_{12}+\pi_{44}}{2}$	$[\bar{1}10]$	$\frac{\pi_{11}+\pi_{12}-\pi_{44}}{2}$
(110)	[111]	$\frac{\pi_{11}+2\pi_{12}+2\pi_{44}}{3}$	$[\bar{1}12]$	$\frac{\pi_{11}+2\pi_{12}-\pi_{44}}{3}$
	[110]	$\frac{\pi_{11}+\pi_{12}+\pi_{44}}{2}$	[001]	$\frac{\pi_{11}+5\pi_{12}-\pi_{44}}{6}$

with glass-counter body for absolute pressure measurement with an edge length of 650 μm (fig. 10.15(a)). In case of a suitable packaging these sensors can be arranged in an array, to measure the uniaxial force- or pressure-distribution on a surface.

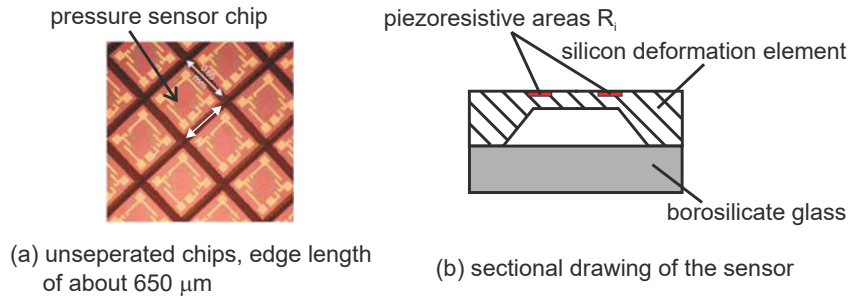
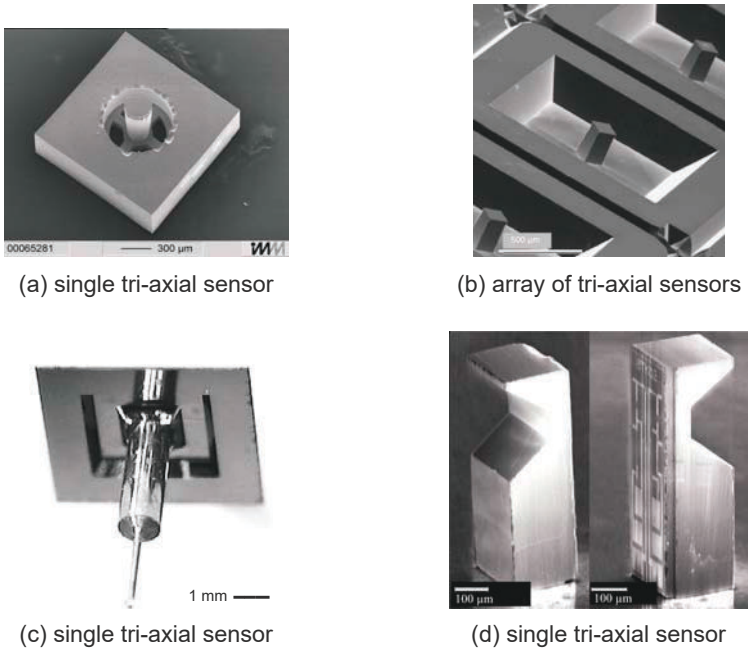


Fig. 10.15 Example of piezoresistive silicon pressure sensors [235].

In case of force sensors the realization of miniaturized multi-component force sensors is subject to actual research. The dimensions of single sensor elements range from 200 μm to 2 mm. The nominal force covers a range of 300 mN to 2 N. Due to the batch-manufacture of measurement elements, the realization of both single sensor elements and array-design[5] is possible. The sensitivity of the sensors takes values of 2 % relative resistivity change in loading case. Figure 10.16 shows four examples of actual topics of research. Variants (a) [266], (b) [271] and (d) [171] were designed for the force measurement in haptic systems. Variant (c) [27] was built for tactile, dimensional measurement technology. The force transmission is always realized by beam- or rod-like structures.

Since 2007 a Hungarian manufacturer is selling the *Tactologic* system. Up to 64 miniaturized sensor elements are connected in an array of 3x3 mm^2. The sensor elements have a size of 0.3x0.3 mm^2 and are able to measure shear forces up to 1 N

[5] By isolating arrays instead of single sensors in the last processing step.

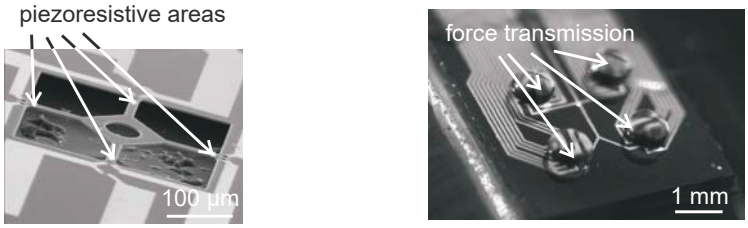


(a) single tri-axial sensor (b) array of tri-axial sensors

(c) single tri-axial sensor (d) single tri-axial sensor

Fig. 10.16 Different realisations of piezoresistive silicon force sensors: [266] (a), [271] (b), [27] (c), [171] (d).

and normal force up to 2.5 N at nominal load. The force transmission is realized by soft silicone dots, applied to every individual sensor element (fig. 10.17 (a) and (b)). Using this array, static and dynamic loads in the range of kilohertz are measureable. But the viscoelastic material properties of the force transmission influence the dynamics due to creeping, especially the measurement of the normal forces [270], [269].



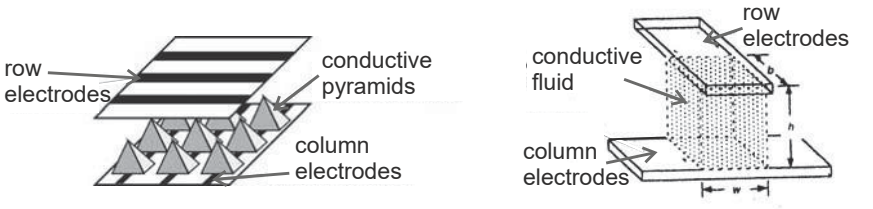
(a) one taxel of the sensing array (Tactologic)

(b) 2x2 array for triaxial-force measurement (Tactologic)

Fig. 10.17 Tactile multi-component force sensor [269].

10.2.4 Further Resistive Sensors

Besides the resistive transducers presented until now, other more "exotic" design exists, which will be introduced with three examples. All sensors are suitable for array assembly to measure position-dependent pressure and a single force component. The used measurement principles are based on the change of geometrical parameters of the force elements. The examples shown in figure 10.18 (a) [139] and (b) [212] use the load-dependency of the constriction resistance. With increased pressure[6] the electrical contact area A increases and the resistance decreases. The companies Interlink and TekScan use this effect for their sensor arrays (also called *Force Sensing Resistors - FSR*). Interlink distributes polymer-foils printed with resistor pastes in thick-film technology. Their basic resistance takes values in the region of $M\Omega$. The sensor foils have a height of 0.25 mm and a working range of zero to one Newton respectively 100 N. Beside the sensitivity to force or pressure, the sensors show a temperature dependency of 0.5 %K. The sensor foils from TexScan are located in the range of 4.4 N up to 440 N, the spatial distribution reaches up to 27.6 elements per centimetre. The available array size reach from approximately 13x13 mm^2 up to 0.5x0.5 m^2. The height of the foils is around 0.1 mm. The measurement inaccuracy takes a value of 10 %. The frequency range reaches from static up to 100 Hz. Beside the application in data gloves, as described by BURDEA [34], the foil sensors are used in orthopedy to detect the pressure distribution in shoes and prosthesis and within automotive industry for ergonomic studies. Another approach is the variation of the distance between two electrodes (Variant (b) in figure 10.18). The sensing element is made of flexible substrates. The electrodes are arranged in rows and columns. The gaps in between are filled with an electrically conductive fluid. In loading case the fluid is squeezed out and the distance of the electrodes varies. A disadvantage of this principle is given by the necessity for very large distance variations up to 10 mm to achieve usable output signals. Until today this principle is still a topic of research.



(a) micromachined tactile array developed by Fraunhofer Institute IBMT

(b) variation of the electrodes distance

Fig. 10.18 Selected examples of foil sensors using the effect of a load-dependent constriction resistance [139], [212].

[6] The force can be calculated taking the contact area into account.

10.2.5 Capacitive Sensors

Within every capacitive sensor at least two electrodes are parallel located to each other. Figure 10.19 shows a design based on a single measurement capacity.

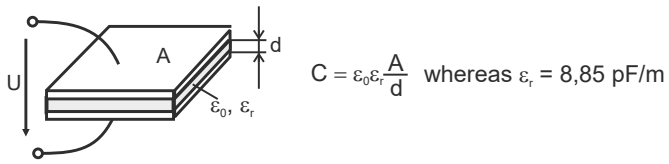


Fig. 10.19 Assembly of a single capacitance.

In contrast to the resistive principle - measuring the mechanical variables stress and strain - the capacitive principle measures the integral values displacement (or elongation) directly.

Concerning the working principle, three classes can be identified, which shows some similarities to electrostatic actuators discussed in chapter 9.5. The first class uses the displacement principle. On this, the mechanical load changes the electrode distance d or the active electrode area A. In the third class the relative dielectric ε_r is influenced. The change of electrode distance is usually used for measuring force, pressure, displacement, and acceleration. In these cases the mechanical load is directly applied to the electrode and displaces it relatively to the other one. The resulting capacitance change can be calculated:

$$\frac{\Delta C}{C_0} = \frac{1}{1 \pm \xi / d} \approx \pm \frac{\xi}{d}. \tag{10.18}$$

ξ marks the change of distance. Additionally, the electrode distance can be kept constant, and only one electrode can be parallel displaced (fig. 10.20). The active electrode area varies accordingly and the resulting capacitance change can be used to measure angle, filling level, or displacement. It is calculated according to:

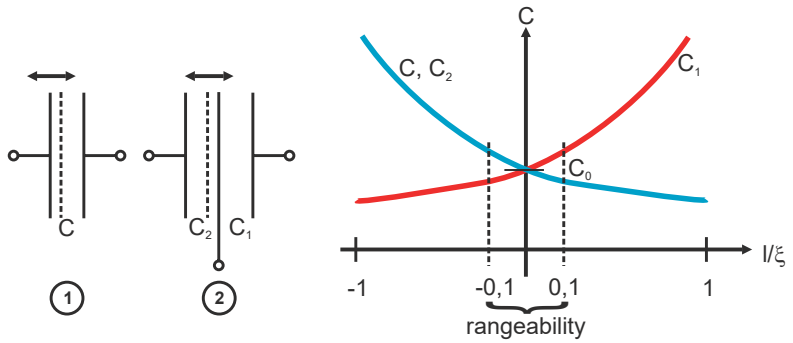
$$\frac{\Delta C}{C_0} = 1 \pm \frac{\Delta A}{A_0}. \tag{10.19}$$

The third option for a capacitance change is the variation of the relative dielectric. This principal is often used for measuring a filling level e.g. of liquids, or as a proximity switch for layer thickness. This capacitance change is calculated according to

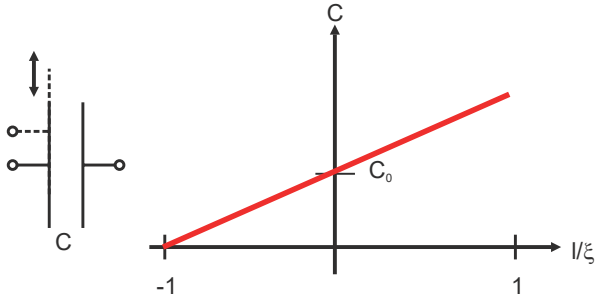
$$\frac{\Delta C}{C_0} = 1 \pm \frac{\Delta \varepsilon_r}{\varepsilon_{r0}}. \tag{10.20}$$

10.2.5.1 Characteristics of Capacitive Pressure and Force Sensors

The main principle used for capacitive force- respectively pressure transducers is measuring displacements. Consequently, the following paragraph will concentrate on this principle. As stated within equation 10.18 for the change of distance, the interconnection between capacitance change and mechanical load is nonlinear for single capacities. The displacement ξ lies in the range of 10 nm to 50 μm [284]. For linearization of the characteristic curves an operating point has to be found, e.g. by arranging three electrodes as a differential capacitor. The displacements ξ typical for the working range are $\leq$ 10% than the absolute electrode distance d. In this range the characteristic curve can be approximated as linear (fig. 10.20(a)). With the principle varying the electrode's surface the capacitance changes proportional to it, resulting in a linear capacitance change (fig. 10.20(b)).



(a) variation of electrode distance in case of single (1) and differential setup (2)



(b) variation of effective electrode area

Fig. 10.20 Schematic view of capacitive sensors and their characteristic curve of capacitance.

The evaluation of the capacitance change can be made by an open- or closed-loop measuring method. Concerning open-loop method, the sensor is either integrated in a capacitive bridge circuit, or it is put into a passive oscillating circuit with coil

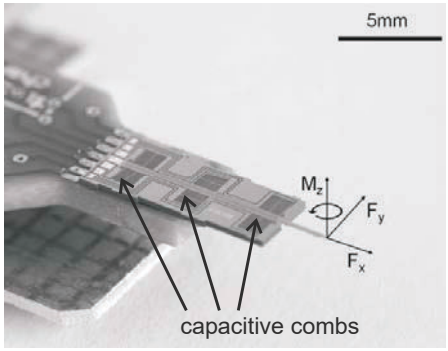
and resistor. Alternatively, the impedance can be measured at a constant measurement frequency. An alternative could be the application of a pulse-width-modulation (also called: re-charging method). A closed-loop approach is characterized by the compensation of the displacement by an additional energy. The main advantage of the closed-loop signal conditioning is the high linearity achieved by very small displacements. Additional information can be found in [284] and [285]. The advantage of capacitive sensor in contrast to resistive sensors lies in the effect of little power consumption and high sensitivity. Additionally, the simple structure enables a low-cost realization of miniaturized structures in surface micro-machining (fig. 10.21). In contrast to the resistive sensors - where positions and dimensions of the resistive areas have a direct influence on transfer characteristics - the manufacture tolerances for capacitive sensors are quite high. Mechanically induced stress due to packaging and temperature influence has almost no influence on their performance. Even miss-positioning of electrodes next to each other do not change the transfer characteristics, only the basic capacitance. The manufacturing steps with silicon capacitive sensors are compatible to CMOS-technology. This allows a direct integration of the sensor electronics on the chip, to minimize parasitic capacities. Especially with miniaturized sensors[7] a good signal-to-noise ration can be achieved [201]. The problem of parasitic capacities or leakage fields is one of the major challenges for the design of capacitive actuators, as it achieves easily a level comparable to the capacitance used for measurement. An additional challenge can be found in the constancy of the dielectric value, which is subject to changes in open air-gap actuators due to humidity or other external influence factors.

10.2.5.2 Examples of Capacitive Sensors

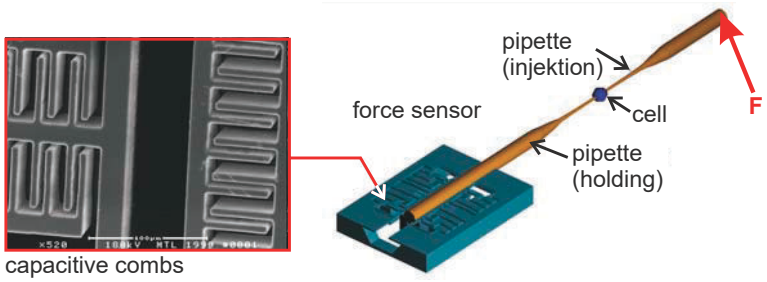
Concerning the manufacturing technology capacitive sensors integrated in haptic systems can be distinguished in three classes. The first class is represented by miniaturized pressure sensors, being realized using silicon microtechnology. Due to their small size of few millimetres the moving masses of the sensor are low and thus cover a wide dynamic range (frequencies from static to several kilohertz). As shown before, the micro-machined capacitive sensors may be combined to arrays for measuring spatially distributed load. As an example SERGIO [223] reports the realization of a capacitive array in CMOS-technology. A challenge is given by the capacity changes in the range of femto Farad, which is similar to the capacity of the wiring. A relief is given by a parallel circuit of several capacities to a sensor element [284]. The frequency range of the shown examples range from static measurement up to several MHz upper cut-off frequency. Consequently, it is suitable for haptic related measurements of tactile information. Another example is given by an array made of poly-silicon. It has an upper cut-off frequency of 1,000 Hz and a spatial resolution of 0.01 mm^2 suitable for tactile measurements. It was originally designed for acquisitions of finger-prints. REY [207] reports of the use of such an array for intracorporeal

[7] Due to the small electrodes a small basic capacitance is achieved, comp. equation in figure 10.19.

pressure measurement at the tip of a gripper. Once again the leakage capacities are a problem, as they are within the range of the measured capacity changes.



(a) 2-component force sensor, nominal load 1,5 mN



(b) 6-component force/torque sensor, nominal load 500 µN

Fig. 10.21 Examples of capacitive silicon multi-component force sensors [19] and [245].

Two examples of multi-component force sensors built in surface micro-machining are shown in figure 10.21 (a) [19] and (b) [245]. The two-axial sensor[8] is designed for atomic force microscopy. The nominal load of this application lie in a range of μN. The three-axial sensor was designed for micro-manipulation, e.g. in molecular biology, with similar nominal values of several μN. Both sensors are using the displacement change for measurement. The second class is represented by ceramic pressure load cells. They are widely used in automotive industry and industrial process measurement technology. Substrate and measurement diaphragm are typically made of Al_2O_3 ceramics. The electrodes are sputtered on ceramic substrates. Substrate and measurement diaphragm are connected via solder applied in thick-film technology. In contrast to silicon sensors, ceramic sensors are macroscopic and have dimensions in the range of several centimetres. Based on this technology sensors in

[8] with respect to "force" component

differential-, relative-, and absolute-designs with nominal pressures in the range of zero to 200 mbar such as in zero to 60 bar are available (e.g. fig. 10.22, Fa. Endress und Hauser). The frequency range of these sensors is low, upper cut-off frequencies of approximately 10 Hz are achieved. The third class is built from foil sensors, dis-

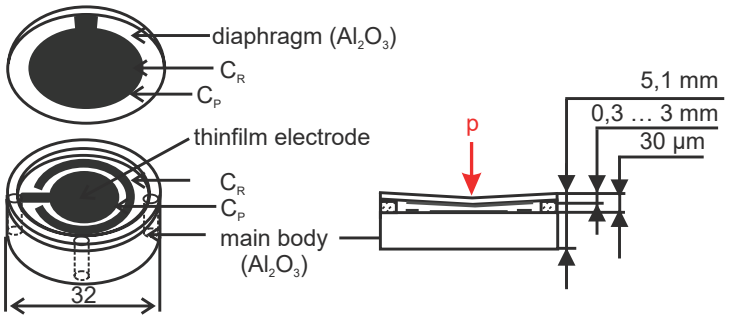


Fig. 10.22 Schematic view of a ceramic pressure sensor fabricated by company Endress und Hauser [284] .

tributed e.g. by the company Althen. These capacitive sensor elements are arranged in a matrix with a spacial resolution of $\leq$ 2x2 mm^2. As substrate a flexible polymer foil is used. The height of such an array is 1 mm. The frequency range ranges from static to approx. 1,000 Hz. Nominal loads up to 200 kPa can be acquired with a resolution of 0.07 kPa. Due to creeping (comp. section 10.1.3) of the substrate and parasitic capacities a high measurement inaccuracy exists. Another polymeric foil sensor in the field of investigation is that one shown in figure 10.23 [38]. In contrast to prior examples this array is used for direct force measurement. Normal forces are detected measuring the change of electrode distance, shear forces by detecting the change of active electrode surface. Similar to the sensor of the company Althen static and dynamic changes up to 1,000 Hz can be measured. The spatial resolution is given with 1x1 mm^2. A disadvantage of the design is the high measurement inaccuracy through creeping of the polymer and leakage capacities.

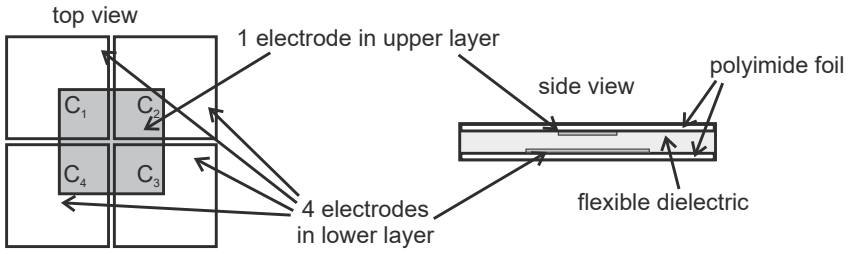


Fig. 10.23 Schematic view of capacitive shear force sensors [38].

10.2.6 Optic Sensors

In the area of optical measurement technology sensors based on freely propagating beams and fiber-optics are available. For force and pressure sensing mainly fiber-optic sensors are used, which will be introduced further within this subsection. All fiber-optic sensors have in common, that the mechanical load influences the transmission characteristics of the optical transmission network, resulting in an influence of the parameters of a reflected or simply transmitted electromagnetic wave. The electromagnetic wave is defined by the wave equation [172].

$$\nabla^2 \Psi = \frac{\delta^2 \Psi}{\delta x^2} + \frac{\delta^2 \Psi}{\delta y^2} + \frac{\delta^2 \Psi}{\delta z^2} \tag{10.21}$$

Ψ represents an arbitrary wave. A possible solution for this differential equation is the propagation of a plane wave in open space. In this case, the electrical field E and the magnetic field B oscillate orthogonal to each other. The electrical field propagating in z-direction is described by equation 10.22.

$$E(z,t) = \frac{1}{2} A(z,t) \cdot e^{j\omega_0 t - \beta_0 z} \tag{10.22}$$

A marks the amplitude of the envelope, ω_0 is the optical carrier frequency and the propagation constant β_0. With the propagation group velocity $v_g(\lambda)$[9] the E- and B-field are connected. Depending on the transmitting medium the group velocity can be calculated via the refraction index n_g [173].

$$v_g(\lambda) = \frac{c_0}{n(\lambda)} \tag{10.23}$$

According to the wave length λ, n different values result. The waves are propagating differently depending on their frequency and wavelength. A pulse "spread out". For further information the sources [172], [173], [190], [298] and [175] are recommended. If only mechanical load such as force or pressure influences the transmission network, the resulting deformation can influence the transmission in two different ways:

1. Material specific: Change of the refraction index n (photoelastic effect)
2. Geometric: Change of beam guidance

The photo-elastic effect describes the anisotropy of the refraction index influenced by mechanical stress. Figure 10.24 visualizes this effect. The resulting refraction index change is dependent on the applied stress T and is given by the following equation [98]:

$$\Delta n = (n_1 - n_2) = C_0 \cdot (T_1 - T_2) \tag{10.24}$$

[9] In vacuum it is equal to the speed of light $c_0 = 2.99792458 \cdot 10^8$ m/s.

C_0 is a material specific, so called photo-elastic coefficient. T_i marks the resulting internal stress. Depending on the refraction index polarization, wave length and phase of beam are changing. In the geometric case, the mechanical load changes the conditions of the beam guidance. Using geometrical optics the influences of mechanical loads on intensity and phase of radiation can be characterized. A disturbing

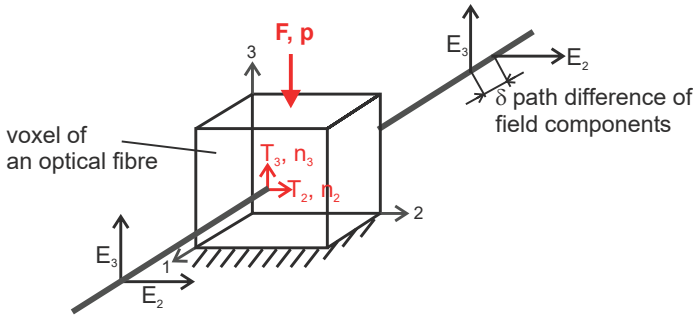


Fig. 10.24 Visualization of the photoelastic effect [98]. Due to the different refraction indices perpendicularly to the propagation direction the propagation velocity of each field component is different and an optical path difference δ occurs. The polarization is changing.

source for all fiber-optical sensors can not be neglected: the temperature. The refraction index is depending on temperature changes, and consequently influences the properties of the guided wave. Beside the thermo-elastic coefficients describing the strain resulting from temperature changes within any material, the temperature directly influences the refraction index itself (paragraph 10.2.6.3). For temperature compensation a reference fiber has to be used, un-loaded and only influenced by temperature change. An advantage of all fiber-optical sensors is given by their immunity to electromagnetic radiation. The following paragraphs introduce the most important principles for optical force- and pressure measurement.

10.2.6.1 Changing Intensity

In principal, two transducer types varying the intensity can be distinguished. Both have in common, that the mechanical load varies the condition of total reflection (fig. 10.25). The angle α_c is defined as the critical angle for total reflexion and defined by SNELLIUS' law:

$$sin(\alpha_c) = \frac{n_2}{n_1} \tag{10.25}$$

The numerical aperture NA gives the appropriate critical angle θ_c for coupling the radiation into a multimode fiber:

$$sin(\theta_c) = \sqrt{n_1^2 - n_2^2} \tag{10.26}$$

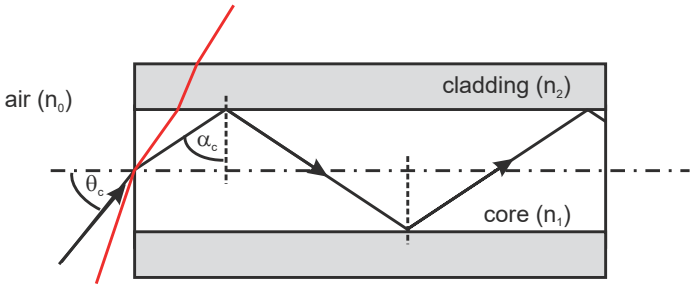


Fig. 10.25 Guidance of multimode fibers. Beams injected with angles above θ_c are not guided in the core.

If the angle varies due to the mechanical load and takes values larger than θ_c resp. smaller values than α_c, the conditions for total reflections are violated. The beam will not be guided within the core of the fiber. The total intensity of the transmitted radiation will become less. Figure 10.26 show a schematic sketch of the design of the very first variant. The sensor element is attached to the end of a multimode fiber. In a first variant the light (e.g. emitted by a laser-diode $\lambda = 1550\,nm$) is coupled

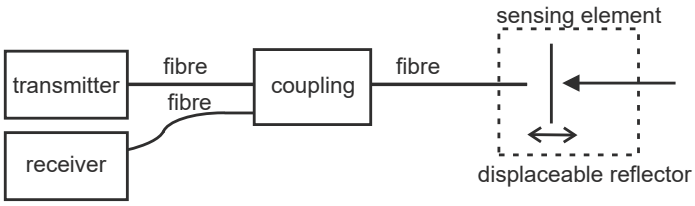


Fig. 10.26 Schematic view of a fiber optic sensor with intensity modulation.

into a multimode fiber. A reflective element is attached to the end of the transmission line. The element can be designed as a deformable object or a rigid one mounted on a deformable substrate. The mechanical load acts on this object. Due to the load the reflective element will be deformed (in case of a flexible surface) or displaced (in case of a rigid surface). Varying the displacement the mode of operation is comparable to a displacement sensor. The intensity is directly proportional to the displacement (fig. 10.27). The load itself is a function of displacement and directly proportional to the elastic compliance n of the sensor element:

$$F(z) = n \cdot z. \tag{10.27}$$

If the geometry of the area changes, a part of the beam - according to the laws of geometrical optics - is decoupled into the cladding (dispersion) and an intensity loss can be measured at the detector (fig. 10.27).

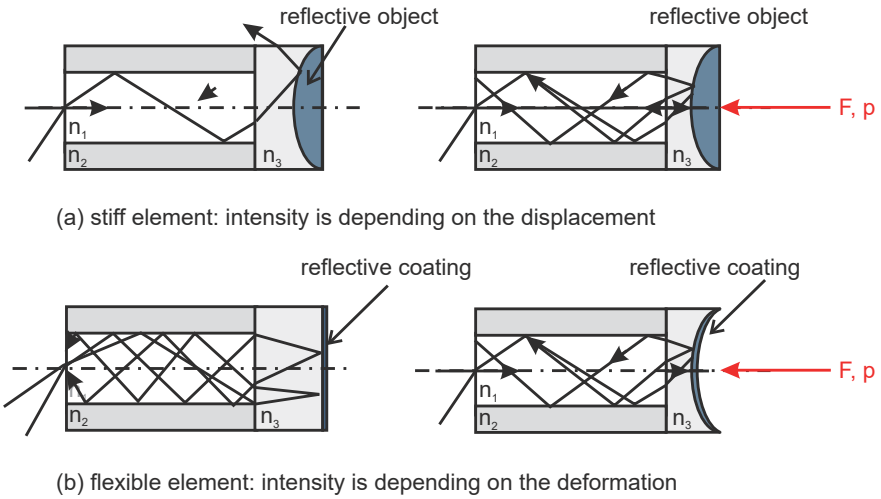


Fig. 10.27 Variation of intensity due to displacement of rigid and flexible elements.

In academically publications from PEIRS [197] and KERN [129] such a mode of operation is suggested for multi-component force measurement. In this case the measurement range is directly proportional to the mechanical properties of fixation of the reflective body. Using the calculation method known from section 10.2.1 this fixation can be designed. A disadvantage of this principal is the use of polymers for the coupling of the reflective object. This leads to creeping of the sensor signal. The measurement inaccuracy of these sensors lies in a range of 10% [129]. Their diameter takes a value of few millimetres. The length depends the application. Another source of noise is the temperature. A temperature change leads to a dilatation (or shrinkage) of the polymer itself and displaces the reflective element. The displacement change results in a defective measurement signal. Due to the small size an array assembly is possible. The second variant is a so called “micro-bending sensor”. Its fundamental design is schematically given in figure 10.28. Like before a beam is coupled into a multimode fiber. Force, pressure or strain is applied by a comb-like structure results in micro-bending of the fiber (fig. 10.28(b)). In case of deformation - similar to the first variant - a part of the light is decoupled into the cladding. The intensity of the measured light diminishes[10]. The gaps between the comb-like structures for micro-bending sensors are in the range of one millimeter. The height of the structure is in the same dimension [243]. To applicate mechanical loads an area

[10] Both versions are possible: Measuring the transmitted and the reflected radiation.

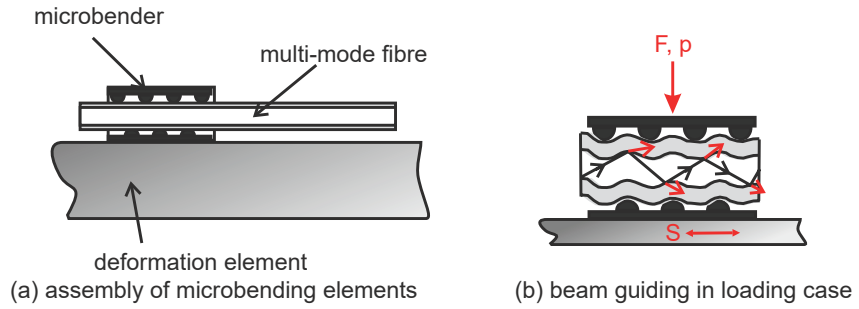


Fig. 10.28 Variation of beam guidance in case of microbending.

of $\cong 1$ cm length and a width of ≥ 5 mm is used. The measurement range depends on the displacement of the bending structure and the diameter of the fiber itself. PANDEY [191] describes the realization of a pressure sensor for loads up to 30 bar. If the bending diameter becomes smaller, lower nominal pressures and forces are possible. Concerning the detection of force components, only one-component sensors can be realized using this principle.

If spatially distributed mechanical load has to be measured, multiple microbending structures can be located along one fiber. To evaluate the several measuring points, for instance optical time domain reflectometry (ODTR) can be used. This device sends a pulsed signal (light pulses of around $10\,\mu$s length) guided in a fiber, and measures the reflexion depending on time. Based on the propagation velocity of the beam inside the fiber v, the time delay for each measuring point can be calculated by relating them. Additional information can be found in [166], [243] or [191]. The dynamics of these sensors is only limited by the sensor electronics and could theoretically be applied to the whole range of haptic applications.

10.2.6.2 Changeing of Phase

The variation of the phase of light by mechanical load is used for interferometric sensors. The most commonly used type is based on the Fabry-Pérot-interferometer, discussed in the following paragraph. Other variants are Michelson- and Mach-Zehnder-interferometers. The assembly is made of two plane-parallel, reflective and semi-transparent objects e.g. at the end of a fiber, building an optical resonator (fig. 10.29). The beam is reflected several times within the resonator and interferes with each reflection. The resonance condition of this assembly is given by the distance d of the reflective elements and the refraction index n within the resonator. The so called "free spectral range" marks the phase difference δ, generating a constructive superposition of beams:

$$\delta = \frac{2\pi}{\lambda} \cdot 2 \cdot n \cdot d \cdot cos(\alpha) \tag{10.28}$$

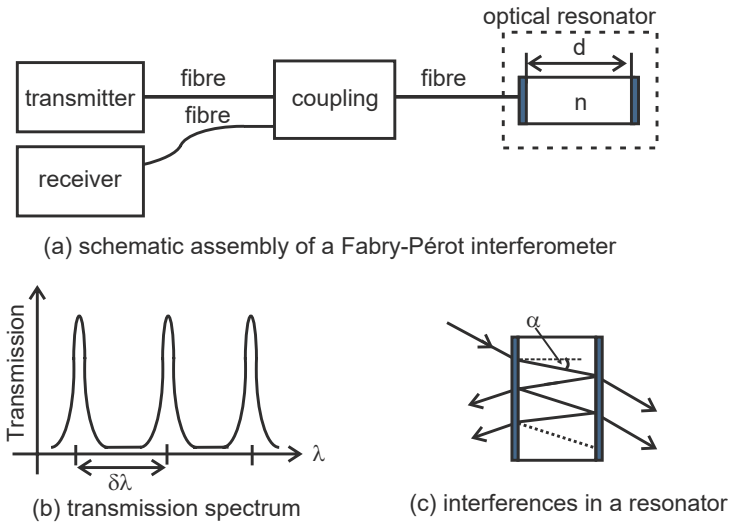


Fig. 10.29 Assembly and operating mode of a Fabry-Pérot interferometer.

Figure 10.29(b) shows the typical characteristics of the transmission spectrum of a Fabry-Pérot interferometer. According to the formula shown above the corresponding wavelength yields a transmission peak; all other wavelengths are damped and annihilated. Due to the mechanical load the distance d of the surfaces is varied, changing the conditions for constructive interference. Sensors using this principle are e.g. used by the company LaserComponents for uniaxial force- or pressure measurement, and can be bought for nominal pressures up to 69 bar [155]. The influence of temperature would also appear to be problematic too and has to be compensated by a reference configuration parallel working. Beside pressure transducers, single component forces and strains can be measured (fig. 10.30). The design equals a Michelson-Interferometer. The sensor element is made of two multimode fibers, whereas the strain acts upon only one fiber. Identical to the Fabry-Pérot-configuration the sensor element is made of two plane-parallel reflective surfaces, whose distance varies according to varying strain. Inside the measuring electronics a reference design is included. To measure the mechanical load the phase of reference and measuring assembly is compared. This measurement principle enables to measure frequencies in the range of several kilohertz. The geometrical dimension is given by the diameter of the fiber including some protective coating ≤ 1 mm, and the length of 2 to 20 mm depending on the application itself. For pressure sensors the measuring error with respect to nominal load takes a value of about 0,5%, with strain gauges at a factor of $15 \cdot 10^{-6}$.

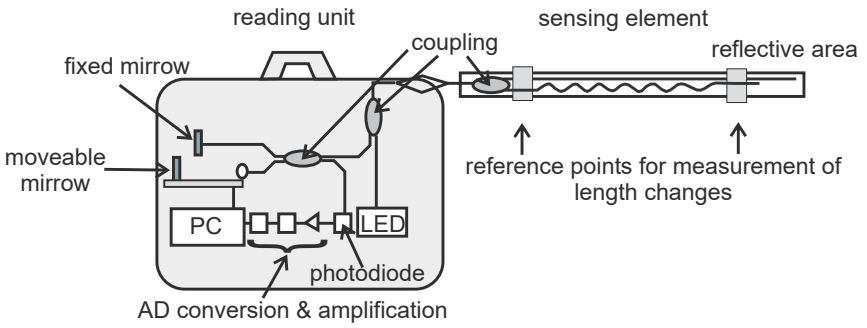


Fig. 10.30 Temperature compensation in interferometric strain sensing elements [108].

10.2.6.3 Changing Wavelength

For optical detection of strain so called fiber BRAGG grating sensors (fbg sensor) are widely used. To realize the sensing element the refractive index of the core in a single mode fiber is varied due to the position (fig. 10.31) and a grating arise [221]. The refractive index modulation can be described by

$$n(z) = n_0 + \delta n_{effective}(z) = n_0 + \delta \overline{n}_{effective} \cdot \left(1 + s \cdot cos\left(\frac{2\pi}{\Lambda} z + \phi(z)\right)\right) \quad (10.29)$$

whereas n_0 is the refractive index within the core, $\delta \overline{n}_{effective}$ is the average of the index's modulation and s a measure of the intensity of the index's modulation. Λ marks the grating period and the phase shift $\phi(z)$ resulting from the measured value. In idle situation results $\phi(z) = 0$. Figure 10.31 gives a schematic drawing of the assembly.

If light is coupled into the fiber, only parts of it are reflected according to the law of BRAGG. The reflective spectrum shows a peak at the so called BRAGG-wavelength λ_b. This wavelength depends on the refractive index $n(z)$ and the grating period Λ:

$$\lambda_b = 2n\Lambda\,. \quad (10.30)$$

In loading case both grating distance and refractive index varies. The maximum of the spectrum is shifting from λ_0 to another wavelength. According to the wavelength shift mechanical load can be determined. The following condition is achieved:

$$\frac{\Delta\lambda}{\lambda_0} = \underbrace{(1 - C_0)}_{k-factor} \cdot (S + \alpha_{VK} \cdot \Delta\vartheta) + \frac{\delta n/n}{\delta\vartheta} \cdot \Delta\vartheta, \quad (10.31)$$

whereas α_{VK} is the coefficient of thermal expansion of the deformation body, and C_0 the photoelastic coefficient. Beside the change induced by mechanical strain S the change of temperature ϑ influences the wavelength shift in the same dimen-

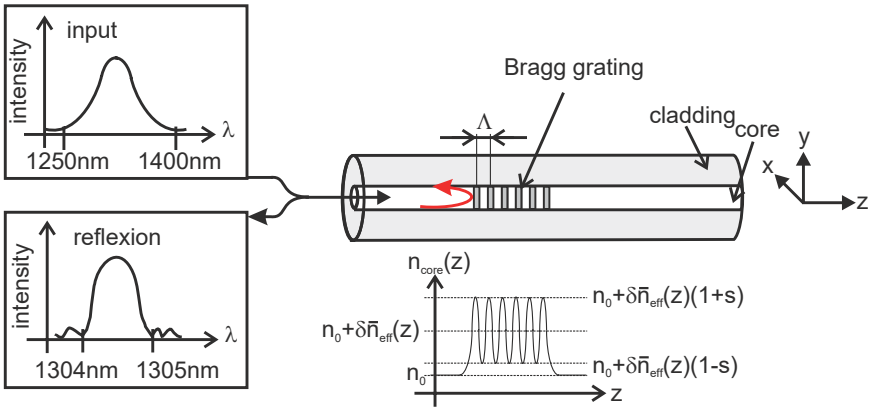


Fig. 10.31 Operational mode of fbg sensors [221] .

sion. Compensating the influence of temperature another FBG sensor has to be installed as reference at an un-loaded area. The temperature compensation is afterwards achieved by comparison between both signals. Analog to resistive strain sensors a k-factor of $k \approx 0.78$ can be achieved with constant measurement temperature. Extensions up to $10.000\,\mu$m/m can be achieved. The width of the sensor lies in the area of single mode fibers. The sensors length is defined by the grating, which has to be three milimetres at least to provide a useable reflective spectrum [133, 26, 53]. The resolution takes a value of $0.1\,\mu$m/m and is - such as its dynamics - defined by the sensor electronics. Similar to strain gauges these sensors can be mechanically applied on deformation elements, whose dimensions and shapes define the measurement range. A challenge with the application of fiber-sensors in this context is the differing coefficients of thermal expansion between deformable element, glue and fiber. Additionally, the reproducibility of the glue-process for fibers is not as high as typically required. Especially creeping of glue results in large measurement errors. Comparable with the micro-bending principle, fbg sensors are applicable to several spatially distributed measurement points. To distinguish the several positions gratings with different periods Λ_i and thus different BRAGG-wavelengths λ_b are used. Further information on the application of fbgs can be found in [221], [53], [26] and [133] .

10.2.7 Piezoelectric Sensors

Piezoelectric sensors are widely used, especially for measurement of highly dynamic activities. The measurement principle is based on a measure induced charge displacement within the piezoelectric material, the so called reciprocal piezoelectric effect (section 9.4). The charge displacement leads to an additional polarization of

the material resulting in a change of charge on the surface. This can be measured using electrodes (fig. 10.32). Beside the measurement of force, it is for pressure- and acceleration measurement especially. For force measurement the longitudinal effect is primarily used. Detailed information about the piezoelectric effect and possible designs are found in section 9.4. The materials uesd for sensing elements will be introduced in the following paragraph.

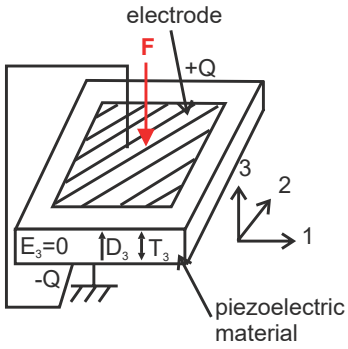
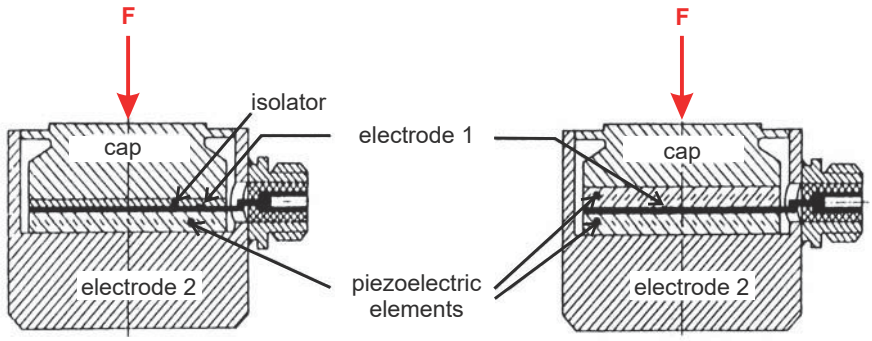


Fig. 10.32 Visualization of piezoelectricity [284].

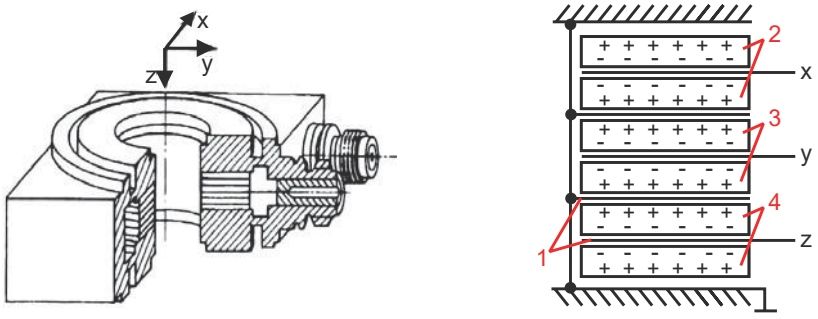
The technically relevant materials can be distinguished into three groups. The first group is built of mono-crystals such as quartz, gallium-, and orthophosphate[11]. The polarisation change in case of mechanical load is direct proportional to the stress. The transfer characteristic is very linear and does not have any relevant hysteresis. The piezoelectric coefficients are long-term stable. One disadvantage is the small coupling factor k of about 0.1. For remembrance: k is defined as the quotient of the transformed to the absorbed energy. The second group is formed by polycrystalline piezo-ceramics, such as barium titanate ($BaTiO_3$) or lead zirconate titanate (PZT, $Pb(ZiTi)O_3$), being manufactured in a sintering process. The polarization is artificially generated during the manufacturing process (section 9.4). An advantage of this material is the coupling factor, which is seven times higher than that one of quartz. A disadvantage is the nonlinear transfer characteristics with a noticeable hysteresis, and a reduced long-term stability. The materials tend to depolarize. The last group is build from partial crystalline plastic foils made of polyvinylidene fluoride (PVDF). Its coupling factor lies with 0.1 to 0.2 in the area of quartz. Advantageous are the limit size (foil thickness of a few μm) and the high elasticity of the material. The first two sensor materials are used in conventional force sensors, as e.g. distributed by the company Kistler. Nominal forces take values of 50 N to 1.2 MN. The sonsor typically has a diameter of 16 mm and a height of 8 mm. Alternations of load up to 100 kHz are measureable. Single- as well as multiple-component sensors are

[11] This crystal is especially applicable for high temperature requirements.

state of the art. 10.33 shows the general design of a three-component-force-sensor from Kistler.



(a) possible assemblies für piezoelectric force sensors



(b) assembly of a 3-component sensor:1 electrodes, 2 quartz plates shear effect - F_x, 3 quartz plates longitudinal effect - F_z, 4 quartz plates schear effect - F_y

Fig. 10.33 Possible assemblies of piezoelectric force sensors [284].

Piezoelectric force sensors are typically used for the analysis of the dynamic forces occurring during drilling and milling or for stress analysis in automotive industry. In haptic system these sensor variants can hardly be found. Not exclusively but mostly because they are not suitable to measure static loads. Sensors based on PVDF-foils as piezoelectric material are increasingly used for the measurement of tactile actions. The piezoelectric effect however is used for the generation of a displacement and not for its measurement, making this variant being described in subsection 10.2.8.1.

10.2.8 Exotic Ones

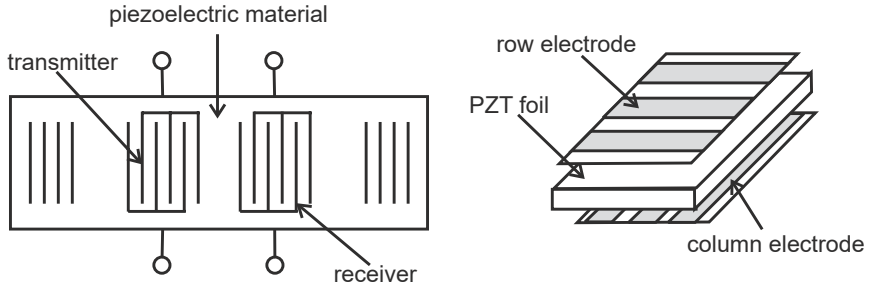
The sensor designs shown in this subsection are not force or pressure sensors for conventional purposes. All of them have been designed for different research projects in the context of haptic systems. The focus of these developments lies in the spatially distributed measurement of tactile information.

10.2.8.1 Resonance Sensors

For measurement of vibrotactile information e.g. the so called resonance principle could be used. Figure 10.34(a) shows the principal design of such a sensor. A piezoelectric foil (PZT or even PVDF) is used as an actuator. Electrodes on both sides of the foil apply an electrical oscillating signal, resulting in mechanical oscillations of the material due to the direct piezoelectric effect. The structure oscillates at its resonance frequency f_0 calculated by the following formula

$$f_0 = \frac{1}{2d} \cdot \sqrt{\frac{n}{\rho}} \tag{10.32}$$

whereas d is the thickness, n the elasticity and ρ the density of the used material. The load, responsible for the deformation, is proportional to the frequency change [200]. For spatially distributed measurement the sensors are connected as arrays of elements with 3x3 and 15x15 sensors. The dimensions of the sensing arrays takes values of 8x8 mm^2 resp. 14x14 mm^2. The thickness of the foil is $<<$ 1 mm. A huge disadvantage of this principle is the high temperature dependency of the resonance frequency from the piezoelectric material used. The coefficient lies at 11.5 Hz per 1 °C within a temperature range between 20 to 30 °C [64], [139]. So called surface



(a) principle of a surface acoustic wave resonator

(b) principle of a resonant sensor array

Fig. 10.34 Schematical view of reonance sensors. (a) [284], (b) [200] .

acoustic wave resonators, SAW sensors, make use of the change of their resonance frequency too. The excitation occurs via an emitter called “Inter-digital structure” (fig. 10.34(b)). The mechanical oscillations with frequencies in the range of MHz distribute along the surface of the material. They are reflected on parallel metal structures and detected by the receiving structure. Due to mechanical values applied the material is deformed, the runtime of the mechanical wave changes, and consequently the sensor’s resonance frequency. With this design, the temperature is one of the major disturbing values. SAW sensors are used for measurement of force, torques, pressure and strain. The dynamic range reaches from static to highly dynamic loads.

10.2.8.2 Electrodynamic Sensor Systems

Within the research project *TAMIC* an active sensor system for the analysis of organic tissue in minimally invasive surgery was developed [216]. The underlying principle is based on an electrodynamic actuated plunger excited to oscillations (comp. section 9.2). The plunger is magnetized in axial direction. The movements of the plunger induce voltages within an additional coil included in the system. The material to be measured is damping the movement, which can be detected and quantified by the induced voltage. The maximum displacement of the plunger is set to one milimetre. The system is able to measure dynamically from 10 Hz to 60 Hz. The nominal force lies in the range of 200 mN. The geometrical dimensions of the system are a diameter of $\leq$15 mm, and a length of $\leq$400 mm, which is near to typical minimally invasive instruments. Detailed information can be found in [222].

Another example for a miniaturized sensor for the measurement of spatially distributed tactile information is presented by HASEGAWA in [86]. Figure 10.35 shows the schematic design of one element. The elements are arranged in an array struc-

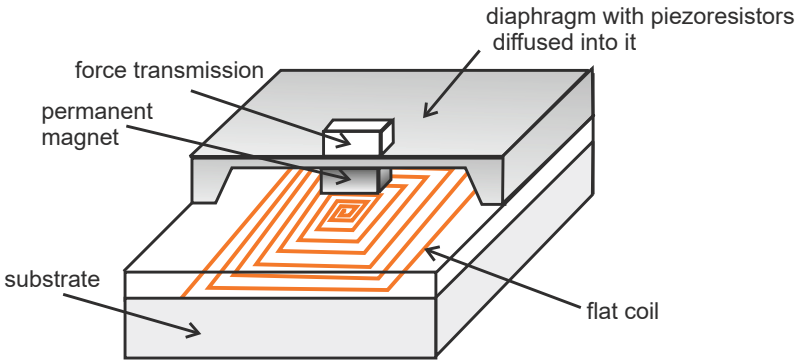


Fig. 10.35 Schematic view of an active element [86]. The dimensions are 6x6x1 mm^3.

ture. In quasi-static operation mode the system is able to measure contact force and the measurement object's elasticity. The upper surface is made of a silicon-diaphragm with a small cubical for force-application to the center of the plane. The displacement of the plate is measured identical to a silicon-pressure or -force sensor with piezoresistive areas on the substrate. By the displacement the applied contact force can be derived. For measuring the elastic compliance of the object, a current is applied to the flat coil (fig. 10.35). In the center of the diaphragm's lower side a permanent magnet is mounted. The electrically generated magnetic field is oriented in the opposite direction of the permanent magnet. The plate is displaced by this electromagnetic actuator and the cube is pressed back into the object. The force necessary to deform the object is used in combination with the piezoresistive sensors signal for calculation of the object's elastic compliance. In the dynamic operation mode the coil is supplied with an oscillating signal, operating the diaphragm in resonance. Due to interaction with the measured object the resonance condition changes. By the changing parameters, such as phase rotation, resonance frequency, and amplitude, elastic coefficients such as damping coefficients of the material can be identified. Due to the high degree of miniaturization highly dynamic actions up to several kilohertz are possible to be measured. The nominal force lies in the area of $2\,N$, the resolution of the system is unknown.

10.2.8.3 Electroluminescence Sensors

A high resolution touch-sensors is presented by SARAF [213]. It is thought to be used for the analysis of texture on organ surfaces. Figure 10.36 shows a schematic sketch. On a transparent glass substrate a layer-compound of $10\,\mu$m height made of gold- and cadmium sulfite particles[12] is applied. The single layers are separated by dielectric barriers. The mechanical load is applied on the upper gold layer, resulting in a break-through of the dielectric layer and a current flow. Additionally, energy is released in form of small flashes. This optical signal is detected using a CCD-camera. The signal is directly proportional to the strain distribution generated by the load. The resulting current density is measured and interpreted. The spatial resolution of the design is given with $50\,\mu$m. Nominal pressures of around 0.8 bar can be detected. The sensor area has a size of 2.5x2.5 mm^2, the thickness of the sensor is ≤ 1 mm and thus very thin. Additional information can be found in [207].

10.3 Selection of a Suitable Sensor

In earlier sections first of all sources for the requirements identification have been presented. Afterwards, the presentation and discussion of the most relevant sensor principles to measure forces was made. This section is thought as an aid, to select

[12] a semi-conducting material

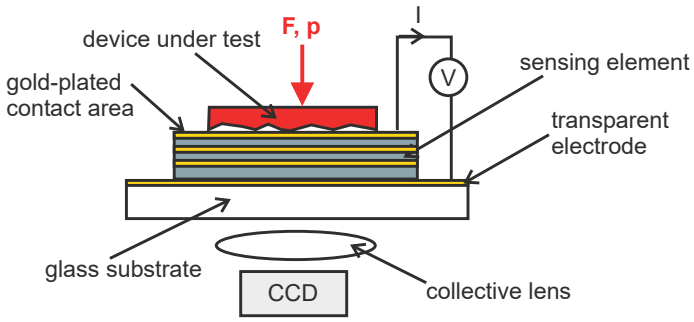


Fig. 10.36 Schematic view of an electroluminiscence sensor [213].

an appropriate force sensor. Depending on the identified requirements found using section 10.1, a suitable sensor principle can be chosen.

To get a better overview, the basic requirements described in section 10.1.5 are collected in table 10.4. The requirements are distinguished concerning the human perception in kinaesthetic and tactile information. More detailed information concerning force- and spatial-resolution can of course be found in section 10.1. The properties of active and passive transformers - the force measurement is done via a mechanical variable such as strain or stress detected with elastomechanics - are strongly dependent on the design of the deformation element. Especially the nominal force, number of components to be measured and the dynamics are directly influenced by the deformation element's design.

Table 10.4 Compilation of main requirements on haptic sensors

type of information	requirements	values
kinaesthetic	nominal load F_N	(5 ... 100) N
	resolution ΔF	5 %F_N
	frequency range	(0 ... 10) Hz
tactile	nominal load F_N	$\leq$ 0.3 N bzw. $\leq$ 4.5 N
	resolution ΔF	5 %F_N
	frequency range	(0 ... 1,000) Hz
	spatial resolution Δx	structural dependent, $\geq$ 1 mm

A comparison with other sensor principles can hardly be done. Consequently, the methods will be compared separately from each other. As an evaluation criterion transfer characteristics and geometrical dimensions are chosen. Figure 10.37 classifies the principles according to k-factor and geometry. If the strain sensing element becomes smaller, the smaller the total force sensor design can be. A direct

result of the smaller size is the minimized mass, providing an increased upper cut-off frequency. If the k-factor of the sensing element is higher, lower absolute value of strain is necessary to get a high output signal. Additionally, the overall design can be made stiffer. This enables to detect smaller nominal forces. Concerning the lower cut-off frequency the strain sensing elements are suitable for measuring static loads. Using piezoresistive and capacitive silicon-sensor an upper cut-off frequency of 10 kHz or more can be measured with high resolution. The other sensor principles

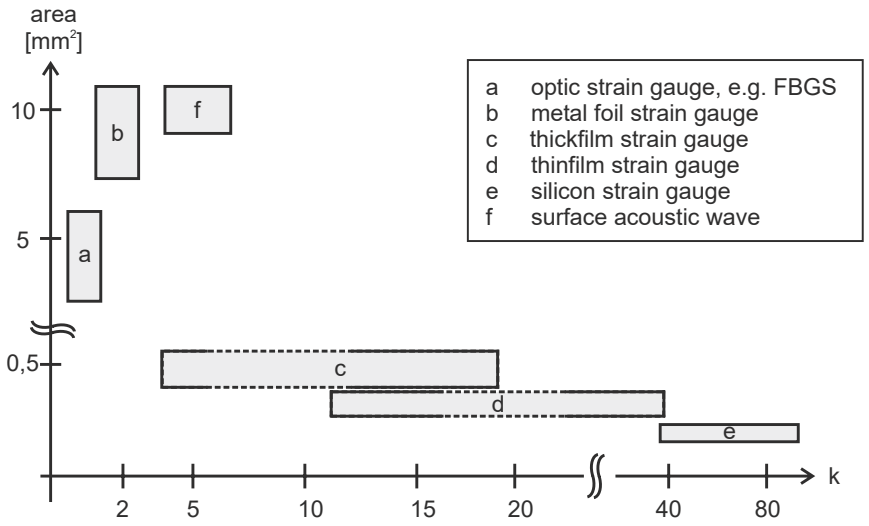


Fig. 10.37 Comparison of different strain measurement technologies due to dimensions and gauge factor.

can be compared contingent on the nominal load and the dimensions. Figure 10.38 classifies the presented principals according to their nominal load and corresponding construction space.

Except the piezoelectric sensors, all sensor principles can be used for measuring static and dynamic loads. The upper cut-off frequency mainly depends on the mass of the sensor which has to be moved. Consequently, the more miniaturized the sensor the higher the upper cut-off frequency becomes. Figure 10.39 compares the presented sensor principles according to the detectable nominal load and the corresponding dynamic range.

With the aid of the shown diagrams a pre-selection of suitable sensor principles for the intended application can be done. Additional sensor properties such as resolution, energy consumption, costs or impact of noise are strongly depending on the individual realization and will not be taken into account here. Advanced descriptions of sensor properties have to be taken from literature highlighted in the corresponding subsections for the individual principle.

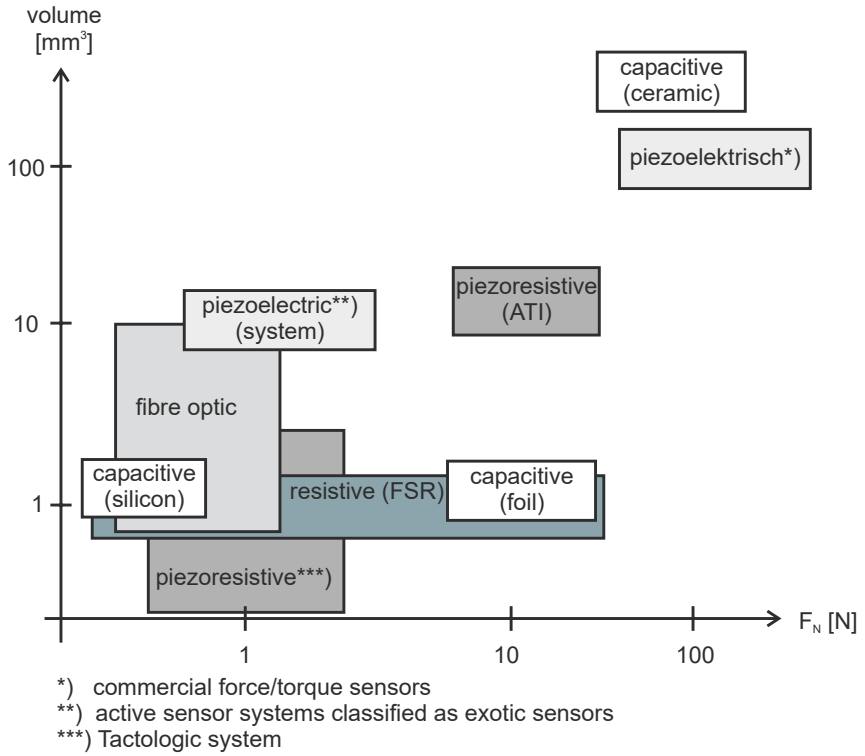


Fig. 10.38 Comparison of different measurement technologies due to dimensions and nominal load.

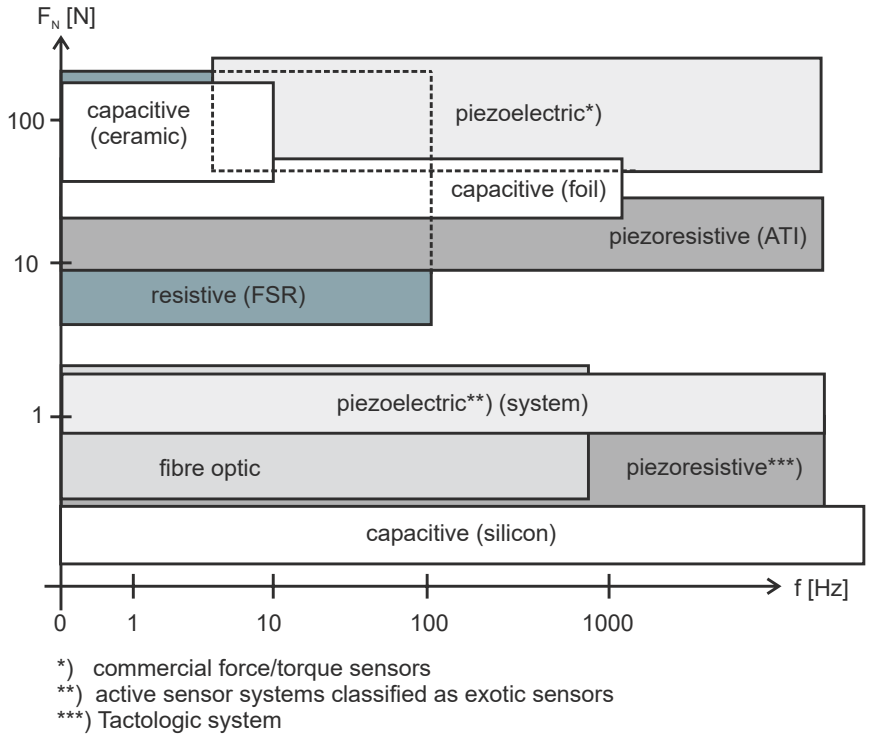


Fig. 10.39 Comparison of different measurement technologies due to nominal load and frequency range.

Chapter 11
Application of Positioning Sensors

THORSTEN A. KERN

To acquire the user's reaction in haptic systems, a measurement of positions respectively their time derivatives (velocities, accelerations) is necessary. Several measurement principles are available to achieve this. A mechanical inlfuence of the sensor on the system has to be avoided for haptic applications, especially kinaesthetic ones. Consequently this discussion focuses on principles which do not affect the mechanical properties significantly. Beside the common optical measurement principles the use of inductive or capacitive sensors is promising especially in combination with actuator design. This chapter gives an overview about the most frequently used principles, amended by hints for their advantages and disadvantages when applied to haptic systems.

11.1 Basic Principles of Position Measurement

For position measurement two principle approaches can be distinguished: differential and absolute measuring systems.

11.1.1 Incremental Principle

Differential systems acquire the change in discrete steps together with the direction of change, and protocol (typically: count) these events. This protocol has to be set back to a reference position by an external signal. If no step loss happens during movement, a prior initialized differential system is able to provide the absolute position as output. If this initializing reference position is set in point which is passed

T.A. Kern (ed.), *Engineering Haptic Devices*, DOI 10.1007/978-3-540-88248-0_11,
© Springer-Verlag Berlin Heidelberg 2009

often, a differential system will be referenced frequently during normal operation. Potential step losses would then affect the time till the next initializing event only.
Measurement of the steps is done via a discrete periodic event, typically encoded in a code disc with grooves or a magnetic rotor. This event is transformed by the sensor in a digital signal, whose frequency is proportional to the velocity of the movement (fig. 11.1a). Some additional directional information is required to be able to measure the absolute position. A typical solution for this purpose is the use of two identical event types with a phase shift (between 1° and 179°, typically 90°). By looking at the status (*high/low*) of these incremental signals (fig. 11.1b) at e.g. the rising edges of the first incremental signal (A), a *low* encodes one movement direction, and a *high* encodes the opposite movement direction. Accordingly the count process either adds or subtracts the pulses generated - in this case - by the second signal (B). State-of-the-art microcontrollers are equipped with counters for incremental measurement already. They provide input pins for count-signal and count-direction. Discrete counters are sold as "Quadrature-Encoder" ICs and frequently include actuator drive electronics, which can be applied for positioning tasks. Latter prevents them from being useful for typical haptic applications.

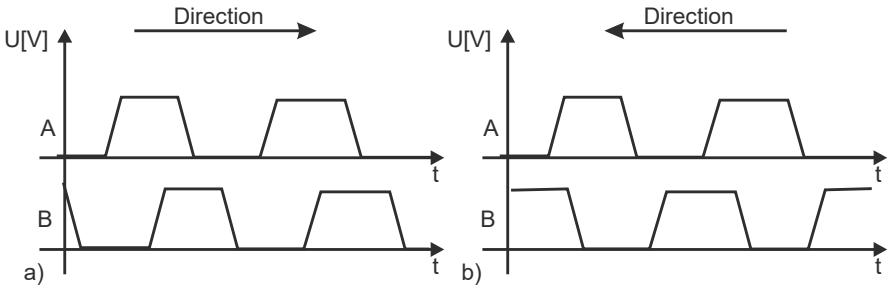


Fig. 11.1 Principle of direction detection with two digital signals with a 90° phase-lag.

11.1.2 Absolute Measurement Principle

Absolute measurement systems acquire a position- or angle-proportional value directly. They are usually analog. A reference position for these systems is not necessary. They have advantages with reference to their measurement frequency, as they are not required to measure with a dynamic defined by the maximum movement velocity. The acquisition dynamics of incremental principles is given by the necessity not to miss any events. In case of absolute measurement principles the measurement frequency can be adjusted to the process-dynamics afterwards, which is usually less demanding. However by the analog measurement technology the efforts are quite high for the circuit, the compensation of disturbances, and the almost obligate digi-

talization of the analog signal.
An alternative for the pure absolute measurement with analog technology is given by a discrete absolute measurement of defined states. In section 9.2.2.1, figure 9.15, a commutation of EC-drives with a discrete, position coding of magnet-angles with field plates was already shown. This approach is based on the assumption to achieve with m measurement points n states in a discrete resolution ΔD.

$$n^m = \Delta D. \tag{11.1}$$

In case of the commutated EC-drive $m = 3$ measurement points, which are able to have $n = 2$ states, could encode 8 positions on the circumference, but only six were actually used. But there are other more complex code discs with several lanes for one sensor each. These sensors are usually able to code two states. However e.g. by the use of different colors on the disc many more states would be imaginable. A resolution of e.g. 1° (360 discrete steps) would need the number of

$$m = \frac{\log(\Delta D)}{\log(n)} = 8.49 \tag{11.2}$$

at least nine lanes for encoding.

11.2 Requirements in the Context of Haptics

Position measurement systems are primarily characterized by their achievable resolution and dynamics. For haptic devices, in dependence on the measurement basis for computer mice and scanners, position resolutions are frequently defined as dots-per-inch ΔR_{inch}. Consequently the resolution ΔR_{mm} in metric millimeters is given as:

$$\Delta R_{\text{mm}} = \frac{25,4\,\text{mm dpi}}{\Delta R_{\text{inch}}}. \tag{11.3}$$

A system with 300 dpi resolution achieves an actual resolution of 84 μm. In dependency on the measurement principle used, different actions have to be taken to achieve this measurement quality. With incremental measurement systems the sensors for the acquisition of single steps (e.g. holes on a mask) are frequently less resolutive, requiring a transformation of the user's movement to larger displacements at the sensor. This is typically achieved by larger diameters of code discs and measurement at their edge. These discs are mounted on an axis, e.g. of an actuator. With analog absolute systems an option for improving the signal is conditioning. It is aimed at reducing the noise component in the signal relative to the wanted signal. This is usually done by a suppression of the noise source (e.g. ambient light), the modulation and filtering of the signal (e.g. lock-in amplifier, compare section 11.6) or the improvement of secondary electronics of the sensors (high resolutive A/D-

transformer, constant reference sources).

Beside the position measurement itself, its dynamic has to be considered during the design process. This requirement is relevant for incremental measurement systems only. Absolute measurement systems need a bandwidth equal to the bandwidth provided by the interface and the transmission chain (chapter 12) for positioning information. Incremental measurement systems however have to be capable of detecting any movement event, independent from the actual movement velocity. The protocol format, usually given by counters part of the microcontrollers, has to be dimensioned to cover the maximum incremental frequency. This requires some assumptions for the maximum movement velocity v_{max}. If e.g. a system with 300 dpi position resolution move at a maximum velocity of 100 mm/s, the dynamic f_{ink} for detecting the increments is given as

$$\frac{1}{f_{ink}} = \frac{\Delta R_{mm}}{v_{max}} \tag{11.4}$$

For the example the necessary measurement frequency is given with $f_{ink} = 1190\,\text{Hz}$. The effective counting frequency is usually chosen with factor two to four higher than that, to have a security margin for counting errors and direction detection.

11.3 Optical Sensors

Optical Sensors for position measuring are gladly and frequently used. They excel by their mechanical robustness and good signal to noise ratios. They are cheap and in case of direct position measurement quite simple to read out[1].

Code Discs

Code discs represent the most frequently used type of position measurement systems with haptic devices, especially within the class of low-cost devices. They are based on transmission (fig. 11.2a) or reflection of an optical radiation, which is interrupted in discrete events. The necessary baffle is located near to the receiver. It is manufactured by stamping, or printed on a transparent substrate (glass, plastic material) via thigh-film technology or laser printers. For high requirements on resolution they are made of metal, either self-supportive or on a substrate again. In theses cases the openings are generated by a photolithographic etching process. The receivers can be realized in different designs. Figure 11.2 shows a discrete design

[1] The examples presented here are discussed either for translatory and rotatory applications. But all principles may be applied to both, as a translation is just a rotation on a circle with infinite diameter.

with two senders in form of diodes and two receivers (photodiode, phototransistor). The placement of sender/receiver-units have to allow the phase shift for directional detection (section 11.1.1). An alternative is given by a fork light barriers already including a compact sender/receiver unit. Additionally opto-encoder (e.g. HLC2705) exists including the signal conditioning for direction-detection from the two incremental signals. The output pins of these elements provide a frequency and one signal for the direction information.

Gray Scale Values

With similar components, but for absolute measurement a gray scale disc or gray scale sensor can be built. Once again there are transmission and reflection (fig. 11.2b) variants of this sensor. In any case the reflection/transmission of the radiation varies dependent on the angle or position of a code disc. The amplitude of the reflection gives absolute position information of the disc. For measurement, once again, either a discrete design or the usage of integrated circuits in form of so called reflection sensors is possible. Although such sensors are frequently used as pure distance switches only, they show very interesting proportional characteristics between the received numbers of photons and their output signal. They are composed of a light emitting diode as sender and a phototransistor as receiver. In some limits the output is typically given by a linear proportional photoelectric current.

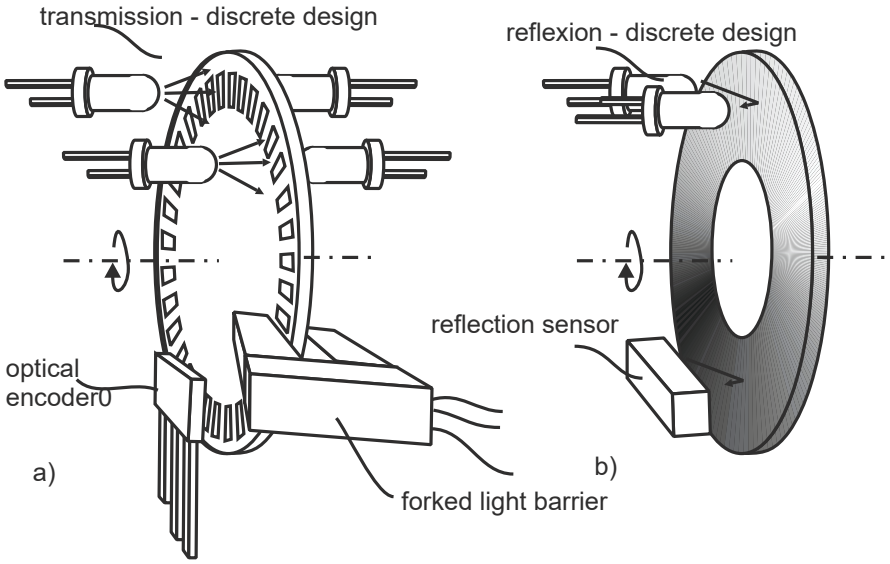


Fig. 11.2 Incremental optical position measurement (a), and absolute position measurement via gray-scale values (b).

Reflection Light Switches

Reflection light switches show useful characteristics for a direct position measurement too. In the range of several millimeters they have a piecewise linear dependency between photocurrent and the distance from the sensor to the reflecting surface. Consequently they are useful as sensor for absolute position measurement of translatory movements (fig. 11.3a). By this method e.g. with the SFH900 or its SMD successor SFH3201 within a near field up to $\approx$1 mm measurement inaccuracies of some micrometers can be achieved. In a more distant field up to 5 mm the sensor is suitable for measurement inaccuracies of $\frac{1}{10}$ mm still.

Mice-Sensor

The invention of optical mice without ball resulted in a new sensor-type interesting for other applications too. The optical mice-sensors are based on an IC measuring an illuminated surface through an optic element (fig. 11.3b). The resolution of the CMOS sensors typically used range from 16x16 to 32x32 pixels. By the image acquired the chip identifies the movement of contrast difference in their direction and velocity. The interface of the calculated values varies from sensor to sensor. The very early types provided an incremental signal for movements in X and Y-direction identical to approaches with code discs described above. They additionally had a serial protocol included to read the complete pixel information. Modern sensors (e.g. ADNB-3532) provide serial protocols for a direct communication with a microcontroller only. This allowed a further miniaturization of the IC and a minimization of the number of contact pins necessary. The resolution of state-of-the-art sensor is in between 500 to 1000 dpi, and is usually sufficient for haptic applications. Only the velocity of position output varies a lot with the sensor types available at the market, and has to be considered carefully for the individual device design. The frequency is usually below 50 Hz. Additionally early sensor designs had some problems with drift and made counting errors, which could be compensated only by frequent referencing.

The sensors are usually sold for computer-mouse-similar applications and corresponding optics. But beside that it is also possible to make measurements of moving surfaces with an adapted optic design at a distance of several centimeters.

Triangulation

Optical triangulation is an additional principle for contactless distance measurement; however it is seldom used for haptic devices. A radiation source, usually a laser, illuminates the surface to be measured, and the reflected radiation is directed on different positions along a sensor array (fig. 11.4). The sensor array may be

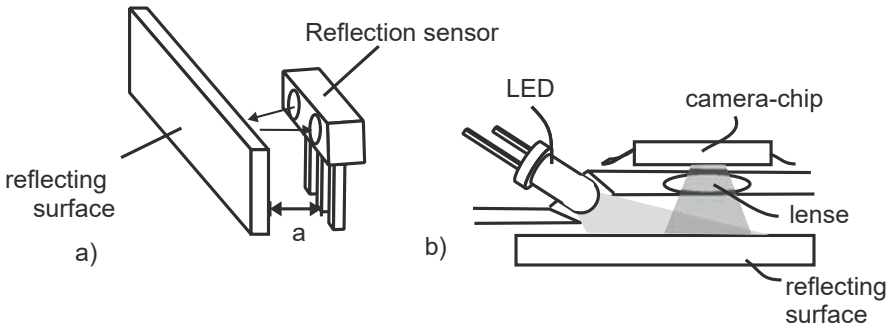


Fig. 11.3 Distance-measurement with reflection light switches (a), and via the movement of an reflective surface in two DOFs "mice-sensor" (b).

made of discrete photodiodes. Frequently it is a CCD or CMOS row with the corresponding high resolution. By focal point identification weighting several detectors a further reduction of measurement inaccuracy can be achieved. Compared to other optical sensors, triangulation sensors are expensive as the detection row with a sufficient resolution is a high cost factor. Their border frequency (»1 kHz) and their measurement inaccuracy (<10 μm) leave nothing to be desired. It is one of the very few principles, which can hardly be used for measuring rotating systems.

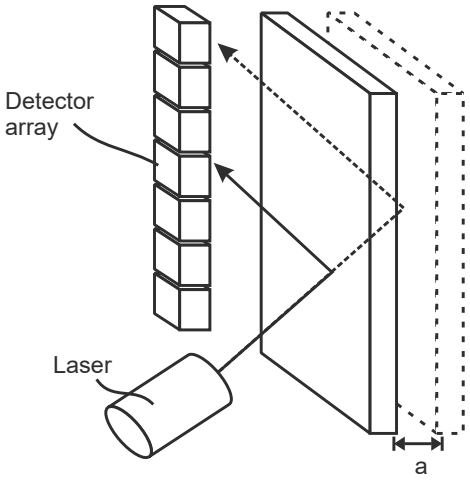


Fig. 11.4 Triangulation of a distance with laser-diode and detector array.

11.4 Magnetic Sensors

Beside the optical measurement principles, especially the group of magnetic measurement principles is relevant for haptic devices. This is a consequence from the fact that electrodynamic and electromagnetic actuators already require magnetic fields to generate forces. For systematization, sensor for static fields, field plates and hall-sensors, and sensor for induced currents and time dependent fields can be distinguished.

Field Plates or Magnetic Dependent Resistors

Field plates or magnetic dependent resistors (MDR) are two pole elements with a resistance being controlled by the presence of a magnetic field. They make use of the GAUSS-effect, which is based on charge carriers being displaced by the LORENTZ-force when crossing a magnetic field. The resulting increase of the path length [187] requires an increase of the ohmic resistance of the material. The parameter characterizing this dependency is dependent on the electron mobility and the path length in the magnetic field. A frequently used material is InSb with very high electron mobility. For an additional increase of the effect, the conductor is formed like in the shape of a meander similar to strain gauges. MDRs are not sensitive to the polarity of the magnetic field. They are detecting the absolute value only. The increase of resistance is nonlinear and similar to a characteristic curve of a diode or transistor. A magnetic pre-load is recommended to use the plates in a linear working point.

Hall-Sensors

Hall-sensors are based on the GAUSS-effect too. In contrast to field plates they are not measuring the resistance increase of the current within the semiconductor, but the voltage orthogonal to the current. This voltage is a direct result of the displacement of the electrodes along the path within the material. The resulting signal is linear and bipolar in dependency on the field-direction. ICs with an integrated amplifier electronics and digital or analog output signals can be bought of the shelf. A frequent use can be found with sensors being located at a phase angle α with diametral magnetized rotational magnets (fig. 11.5). In this application a rotation and rotation-direction is measured.

Inductance Systems

An often forgotten alternative for position measurement is the measurement of changing inductances. The inductance of a system is dependent on many parameters, for example the magnetic permeability of a material in a coil. Using a dif-

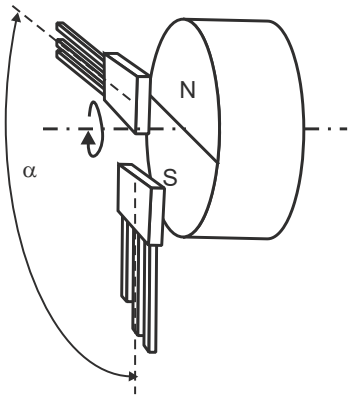


Fig. 11.5 Measurement of the rotation angle of a magnet via field plates or hall-sensors.

ferential measurement in between two coils (fig. 11.6b) a displacement measurement can be made, if a ferromagnetic material moves in between both coils as a position-depending core. As alternatives the geometry of the magnetic circuit may be changed, or its saturation may influence the inductance of the coils. Latter approach is used in systems, where grooves on a ferromagnetic material trigger events in a nearby coil (fig. 11.6a).
A simple electronic for measuring inductance is the use of a LR-serial circuit, which - for example with a microcontroller - is triggered with a voltage step. The measurement value is given by the time the voltage at the resistor needs to trigger a comparator voltage. The duration encodes the inductance, assuming a constant resistance. For the actual design it has to be considered, that the winded coil has an own resistance which cannot be neglected. As an alternative a frequency nearby the resonance $\frac{L}{R}$ of the LR-circuit can be applied. The voltage amplitude measured varies dependent on the inductance's detuned by the movement of the ferromagnetic core.

11.5 Other Displacement Sensors

Beside the displacement measurement principles discussed above there are some rarely used principles still worth to be mentioned here.

Ultrasonic Sensors

Ultrasonic sensors (fig. 11.7) are based on the running time measurement in between the emission of acoustic oscillations and the moment of the acquisition of

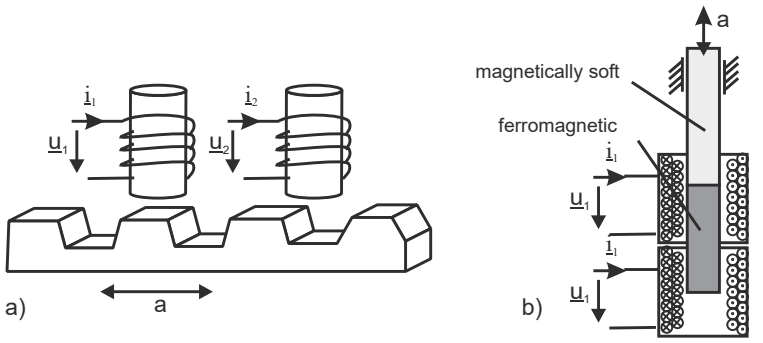


Fig. 11.6 Incremental measurement of a movement via induced currents (a) and differential measurement of the position of a ferromagnetic core (b).

their reflection. The frequency chosen is dependent on the requirements on measurement inaccuracy and the medium for propagation of the wave. As a rough rule of thumb, the denser a material is, the less the damping becomes for acoustic waves. For measurement in tissue frequencies between 1 and 40 MHz are applied. In water frequencies between 100 and 500 kHz and in the atmosphere frequencies well below 30 kHz are used. Whereas with medical applications in tissues the medium shows a damping quite linear in the range of 1dB/Mhz/cm, the measurement within the atmosphere is strongly dependent on the frequency chosen and usually nonlinear. Additionally the acoustic velocity is dependent on the acoustic density of the medium. For the transversal direction - typically used for measurement - velocities between 340 m/s for air and 1500 m/s for water can be achieved. According to the wave theory the minimum measurement inaccuracy possible in transversal direction is $\frac{\lambda}{2}$, which and coupled to both factors mentioned above. It is a natural border of minimum resolution to be achieved.
The most frequently used source and receiver for the mechanical oscillation are piezoelectric materials (fig. 11.7b), whose step response oscillations are sharpened by a coupled mass.

Capacitive Sensors

In section 9.5 the equations for the calculation of capacities between plates of electrostatic actuators (equ. 9.79) were introduced. Of course the measurement of a variable capacity, especially with the linear effect of a transversal plate displacement, can be used for position measurement. This is especially interesting if there are conductive components in the mechanical design, which already move relative to each other. As the capacity is very much dependent on the permittivity of the medium between the plates, which can be strongly influenced by oil or humidity, such a measurement can be done on insusceptible or other well housed actuators

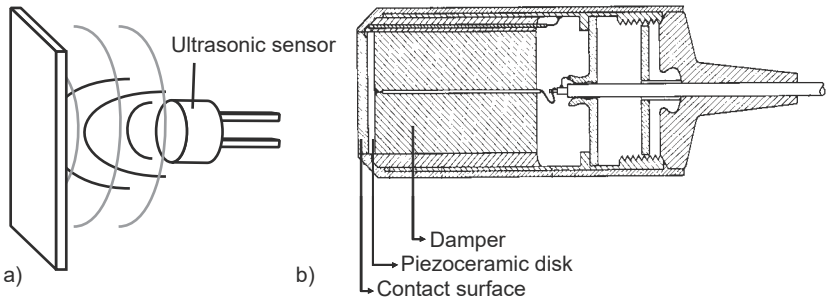


Fig. 11.7 Distance measurement via ultrasonic sensors (a) and cross-section through a medical ultrasonic head with fixed focus (b).

only. Additionally leakage fields of conductors or geometries nearby are usually of the same size as the capacity to be measured. But capacitive sensors for haptic devices can be found in the context of another interesting application. The measurement of the capacity of the handle, even when isolated by a non conductive layer, allows identifying a human touch very securely.

11.6 Electronics for Absolute Positions Sensors

The absolute measurement of a position requires, as mentioned earlier, some additional effort in the electronic design compared to discrete sensors. Two aspects shall be discussed in the context of this chapter.

Constant-Current Supply and Voltage References

For the generation of a constant radiation or the measurement of a bridge circuit the use of constant currents is necessary. There is always the possibility to wire an operational amplifier as a constant current source, or use transistor circuits. Nevertheless for designs with low quantities there are ICs which can be used as current sources directly. The LM234 for example is a voltage-controlled 3-pin IC, providing a current with a maximum error of 0.1%/V change in the supply voltage. The maximum current provided is 10 mA, which is usually sufficient for the supply of optical or resistive sensors.
The change of the signal is usually measured with relation to a voltage in the system. In these cases, it is necessary to provide a voltage which is very well known and independent from temperature effects or changes of the supply voltage. Common voltage regulators as used for electronic supply are not precise enough to fulfil

these requirements. An alternative is given by ZENER-diodes operated in reverse direction. Such diodes however are not applicable to high loads and are of course only available in the steps of ZENER voltages. Alternatively reference voltage sources are available in many voltage steps. The REF02 for example is a six-pin IC, providing a temperature-stable voltage of 5V with an error of 0.3%. The drivable load of such voltage sources is limited, in case of the REF02 it is 10 mA, but this is usually not a relevant limit as they are not thought as a supply to a complex circuit but only as a reference.

Compensation of Noise

The obvious solution for the compensation of noise in a measurement signal is given by the usage of a carrier frequency for modulating the signal. A prerequisite of course is given by the sensor showing no damping at the modulating frequency. This is usually no problem for optical sensors in the range of several kilohertz. At the receiver, the signal is bandpass-filtered and equalized or otherwise averaged. This suppresses disturbing frequencies or otherwise superimposed offsets.
A simple but very effective circuit for noise compensation is the use of so called "lock-in" amplifiers. On the side of the sender a signal is switched between the states *on* and *off* at a frequency f. In the receiver the wanted signal such as e.g. the offset and other disturbing frequencies are receipt. A following amplifier is switched with the same frequency between +1 and -1 in such a way, that with the receipt of the wanted signal including the disturbance the positive amplification happens. During the period without the wanted signal, when the receiver measures the disturbing values only, the signal is inverted with -1. The resulting signal is low-pass filtered afterward, resulting in a subtraction of the noise signal and providing a voltage proportional to the wanted signal only.

11.7 Acceleration and Velocity Measurement

Beside the direct position measurement, haptic systems sometimes demand some knowledge about the first or second derivative of position in form of velocity or acceleration. Such a necessity may be given with stability issues for closed-loop systems or impedance behavior of users or manipulated objects. The acquisition can either be done by direct measurement or by differentiation of the position-signal with digital or analog circuits. Additionally it can be imagined to e.g. measure a velocity and calculate the position by integration. The capabilities of integration and differentiation and their limits, such as typical direct measurement principles, are sketched in this section.

11.7.1 Integration and Differentiation of Signals

The integration and differentiation of signals can either be done analog or digital. Both variants have different advantages and disadvantages.

Analog Differentiation

The basic circuit for an active analog integrator is shown in figure 11.8a. It is a high-pass filter, which already gives hints on the challenges connected with differentiation. The high-pass behavior is limited in its bandwidth. The upper border frequency is given by the resonance frequency $f_R = \frac{1}{2\pi RC}$ and by the bandwidth of the operational amplifier. As these components are sufficiently dynamic for haptic applications this should be no problem in practical realization. Due to the negative feedback however the natural bandwidth limit of the operational amplifier at high frequencies has a phase of 90°, adding to the phase of 90° from the differentiation. This makes the circuit sensitive to become electrically instable and oscillate [255]. This effect can be compensated by a serial resistance with a capacity C, which is identical to a linear amplification with the operational amplifier. This diminishes the phase for high frequencies by 45°, resulting in a phase margin to the instable border condition. Analog differentiation is an adequate method for the derivation of velocities from positioning signals. A double analog differentiation needs a careful design of the corresponding circuit, as a number of capacitive inputs are placed in series. Additionally it should be considered that the amplitude of the operational amplifier is limited by the supply voltage. Accordingly the amplitude's dynamic has to be adjusted to the maximum signal change expected.

Analog Integration

The basic circuit of an active analog integrator is given in figure 11.8b. Analog integration is a reliable method from analog calculation technique, but has limited use for haptic applications. The circuit has an upper border frequency given by the resonance $f_R = \frac{1}{2\pi RC}$, and for a non-ideal operational amplifier it has a lower border frequency too. This is a result of the current I_b at the input of the OP-amplifier charging the capacitor with $U_{\text{in}} = 0\text{V}$ continuously. If C=10 μF and $I_b = 1\,\mu\text{A}$, the voltage increases by 0.1 V per second. Whereas in signal processing applications this can be compensated by high-passes in series, for haptic applications covering a bandwidth from several seconds to 10 kHz this behavior is usually not acceptable.

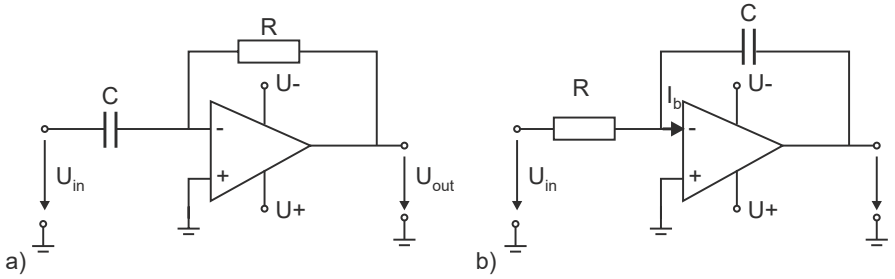


Fig. 11.8 Analog differentiation- (a) resp. integration-circuits (b) [255].

Digital Differentiation

Digital differentiation is realized by a subtraction of two consecutive measurement values. It is very applicable, especially when the signal is measured at high frequencies. The quality of the signal is dependent on the noise on the input. Frequently the least-significant bit of e.g. an AD-conversion is rejected before the differentiation is performed, as it is oscillating with the noise of the AD-conversion (quantization-noise).

Digital Integration

Digital integration is the summation of continuous measurement values and the division of the sum by the number of values. Alternatively it can be the sum of discrete changes of a measurement value. The incremental measurement of a digital encoder is also a form of integration on the basis of change information. The procedure is robust at high frequencies beyond the actual upper border frequency of the signal. Beside a sufficient dimension of the register size for the measurement values to prevent an overflow, there is nothing else to worry about.

11.7.2 Induction as a Velocity Measure

The most frequent variant to gain information about velocity is given by the digital signal processing of a position-measurement. Nevertheless to be able to measure velocity directly, the use of a velocity-proportional physical effect is mandatory. Beside DOPPLER-ultrasonic measurement, which is seldom applicable to haptic systems due to the wavelengths (compare section 11.5), the use of electrical induction is the most frequently used direct effect. Accordingly an electrical induced voltage U is generated in a conductor of the length l, moving orthogonal in a magnetic field B with the velocity v:

$$U = vBl. \tag{11.5}$$

Especially geometrical designs as given with electrodynamic actuators (section 9.2) can be used for velocity measurements with inducing voltages in their coils. In contrast to electrodynamic actuators the design requires a maximization of conductor length, to generate a pronounced voltage signal. The inductivity of the winding generates a low-pass characteristic in combination with its own resistance. This limits the dynamic of the signal. The biggest error made with these kinds of sensors is given by a bad homogeneity of the winding due to dislocation of single turns. This manufacturing error results in different winding lengths moving in the B-field at different positions of the sensor, which is directly affecting the quality of the measured signal.

11.7.3 Force Sensors as Acceleration Sensors

In contrast to velocity measurement, to measure accelerations a wide variety of sensors exists. Ignoring some exceptions, most of them are based on the relation

$$\mathbf{a} = \frac{\mathbf{F}}{m}. \tag{11.6}$$

In fact the force-measurement principles given in chapter 10 are added by a known mass m only, resulting in a mechanical strain of a bending element or generating another acceleration-proportional signal. In professional measurement technology especially piezoelectric sensors for high dynamic measurements, but also piezoresistive sensors for low-frequency accelerations are established. In mechatronic systems with high quantities micromechanic acceleration sensors with comb-like structures in silicon according to the capacitive measurement principle are used. The requirements of automotive industry for airbags and drive stability programs to measure acceleration in many directions made low-price and robust ICs available at the free market, e.g. the ADXL series of Analog Devices. The bandwidth of these sensors ranges from 400 Hz to 2.5 kHz with maximum accelerations >100 g in up to three spatial directions. Only a wide variance of their characteristic values, e.g. the output voltage at 0 g, requires a calibration of the individual sensor.

11.8 Conclusion on Position Measurement

With haptic devices position measurement is a subordinated problem. In the range of physiological perceived displacements resolutions, there are enough sensor principles which are sufficiently precise and dynamic for position measurement. The calculation or measurement of accelerations or velocities is easily possible too. Without

doubt, the optical measurement technology is the most frequently used technical solution. Nevertheless especially for the design of specific actuators it is indicated to ask the questions, whether there are other sensor principles applicable for a direct integration into the actuator. If there are specific requirements for measurement in the range of a few μm positioning resolution, the proposed principles should be treated with reserve. Measurements in the range of μm require specific optical or capacitive measurement technology. With the exception of special psychophysical questions it is unlikely that such requirements are formulated for haptic devices.

Chapter 12
Interface Selection

THORSTEN A. KERN

After the decision for the actuator (chapter 9) used to generate the haptic feedback, and after the measurement of forces (chapter 10) or positions (chapter 11), it becomes necessary to focus on the IT-interface. This interface has to be capable of providing data to the actuation unit and catch and transmit all data from the sensors. Its requirements result - such as with any interface - from the amplitude resolution of the information and the speed at which they have to be transmitted. The focus of this chapter lies on the speed of transmission, as this aspect is the most relevant bottleneck when designing haptic devices. Haptic applications are frequently located on the borderline, may it be with regards on the delay acceptable in the transmission, or the maximum data rate in the sense of a border frequency.

With regards to the interface two typical situations may be distinguished: Spatially distributed tactile displays with a reasonable number of actuators; and primarily kinaesthetic systems with a smaller number of actuators. In case of tactile systems, pin-arrays, vibrators, or tactors the challenge is given by the application of bus-systems for the reduction of cable lengths, and the decentralization of control. Although there are still some questions about timing left, for example to provide tactile signals in the right order despite of a decentralized control, the data rates transmitted are usually not a challenge for common bus systems. VAN ERP points out [267], that a 30 ms time delay between impulses generated by two vibrators at the limbs may not be distinguished any more. For the data interface this observation implies for this application, that any time delay below 30 ms may be uncritical for transmitting information haptically. This is a requirement, which can be fulfilled by serial automation technology network protocols like CAN, or the time triggered version TTCAN, without any problems. Accordingly this section concentrates on requirements of haptic kinaesthetic devices with a small number of actuators only,

T.A. Kern (ed.), *Engineering Haptic Devices*, DOI 10.1007/978-3-540-88248-0_12,
© Springer-Verlag Berlin Heidelberg 2009

whereas these devices usually have to satisfy tactile requirements according to their dynamic responses too.

12.1 Border Frequency of the Transmission Chain

Chapter 4 stated that it is necessary to distinguish two frequency areas when talking about haptic systems. The lower frequency range up to $\approx$ 10 Hz includes a bidirectional information flow, whereas the high frequency area >10 Hz transmits information only unidirectional from the technical system to the user. Although the user himself influences the quality of this transmission by altering the mechanical coupling, this change itself happens at lower frequencies only, and is - from the perspective of bandwidth - not relevant for the transmission. If this knowledge is applied to the typical structures of haptic devices from subsection 2.2.2, some fascinating results can be found. For the following analysis it is assumed, that the transmission and signal conditioning of information happens digital. According to NYQUIST, the maximum signal frequency has to be sampled at least two times faster. In practical application this factor two is a purely theoretical concept, and it is strongly recommend to sample and analog system around 10 times faster than its maximum frequency. The values within figures and texts are based on this assumption.

12.1.1 Bandwidth in a Telemanipulation System

For a telemanipulation system (fig. 12.1) the knowledge about the differing asymmetric dynamics during interaction gives the opportunity to benefit directly for the technical design. In theory it is possible to transmit the haptic information measured at the object within the bandwidth of 1 Hz to 10 kHz, and replay it as forces or positions to the user. The user's reactions may in this case be measured at a bandwidth from static to 5 or 15 Hz only, and be transmitted via controller and manipulator to the object. Although this approach would be functional indeed, the simplicity of position measurement and the necessity to process them for e.g. passivity control result in movements being sampled and transmitted similar dynamic as in the opposite transmission direction for haptic feedback.

12.1.2 Bandwidth in a Simulator-System

For a simulation-system with haptic feedback the different dynamics results in slightly different findings. Nevertheless it is still true, that the movement information may be sampled at a lower rate. However the simulator (fig. 12.2) has to provide the force output at a frequency of 1 to 10 kHz. Due to this simple reason, the simulator

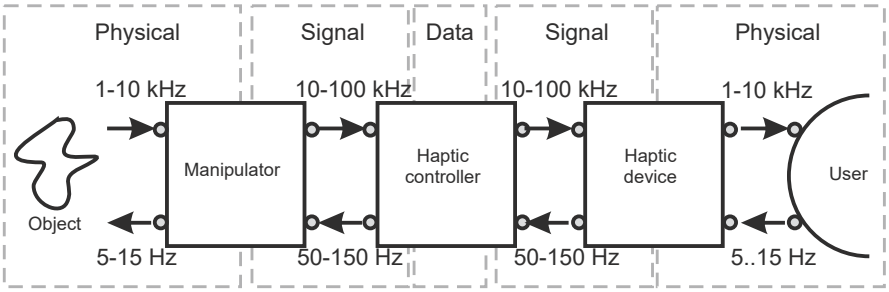


Fig. 12.1 Block diagram of a telemanipulator with haptic feedback.

has to be aware of the actual position data for every simulation step. Consequently with simulators the haptic ouput and the measurement of user reaction has to happen at high frequency (exceptions, see section 12.2).

There are two approaches to integrate the haptic controller in the simulator. In many devices it is designed as an external hardware component (fig. 12.2), which reduces the computing load for the main simulator, and helps reducing the data rate significantly in special data processing concepts with parametrizable models (section 12.2). As an alternative the controller may be realized in software as a driver computed by the simulation main computing unit (fig. 12.3). This is a concept used especially for high power permanently installed simulation machines, or which is used in cost-effective haptic devices for gaming industry with little requirements in dynamics and haptic output.

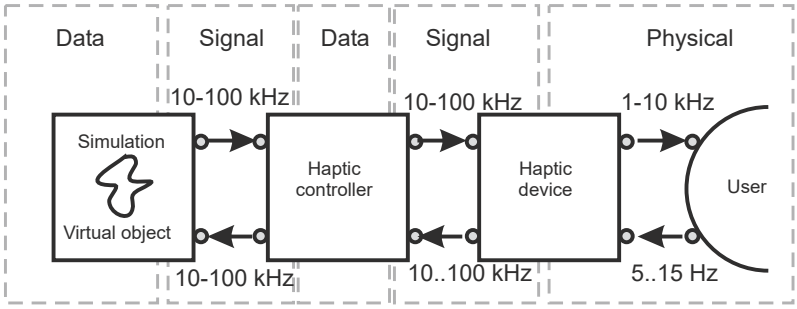


Fig. 12.2 Block diagram of a simulator with haptic feedback and an external controller.

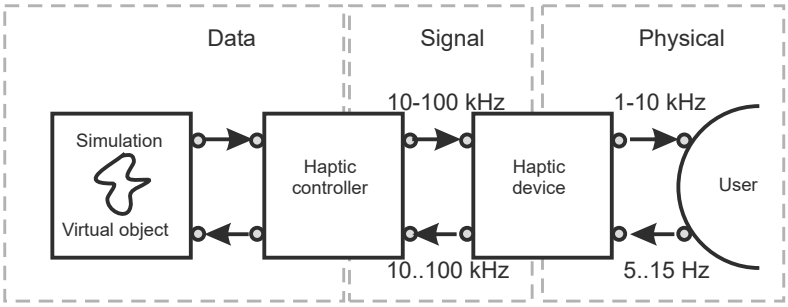


Fig. 12.3 Block-diagram of a simulator with haptic feedback and a controller as part of the driver software .

12.1.3 Data Rates and Latencies

Table 12.1 summarizes the data rates necessary for kinaesthetic applications in some typical examples. The data rates range from 200 kbit/s for simple applications up to 50 Mbit/s for more complex systems. Such rates for the information payload - still excluding the overhead necessary for the protocol and the device control - are achieved by several standard interface types today (section 12.3).

Table 12.1 Example calculating the required unidirectional data rates for typical haptic devices.

DOFs	Resolution	0.1 kHz	1 kHz	10 kHz
1 DOF	8 bit 16 bit	800 bit/s 1600 bit/s	8 kbit/s 16 kbit/s	80 kbit/s 160 kbit/s
3 DOF	8 bit 16 bit	24 kbit/s 48 kbit/s	240 kbit/s 480 kbit/s	2.4 Mbit/s 4.8 Mbit/s
6 DOF	8 bit 16 bit	48 kbit/s 96 kbit/s	480 kbit/s 960 kbit/s	4.8 Mbit/s 9.6 Mbit/s

Beside the requirements for the data rate there is another requirement considering the smallest possible latency. Especially interfaces using packets for transmission, with an uncertainty about the exact time of the transmission (e.g. USB) have to be analyzed critically concerning this effect. Variable latencies between several packets are a problem in any case. If there are constant latencies the reference to other senses with their transmission channel becomes important: A collision is not allowed to happen significantly earlier or later haptically than e.g. visually or acoustically. The range possible for latency is largely dependent on the way to present the other sensual impressions. This interdependencies are subject to current research and are analyzed e.g. by the group around BUSS at the TU-Munich.

12.2 Concepts for Bandwidth Reduction

Whoever ever tried to process a continuous data flow of several megabit with a PC, and in parallel make this PC do some other tasks too, will have noticed that the management of the data flow binds immense computing power. With this problem in mind and as a result from the question about telemanipulation with remotely located systems several solutions for bandwidth reduction of haptic data transmission have been found.

12.2.1 Analysis of the Required Dynamics

The conscious analysis of the dynamics of the situation at hand should be ahead of every method to reduce bandwidth, as already mentioned in chapter 6 and calculated in section 3.3. The limiting cases to be analyzed are given by the initial contact or collision with the objects. If the objects are soft, the border frequencies are in the range <100 Hz. If there are stiff objects part of interaction and if there is the wish to feed back these collisions too, the frequencies up to a border >1 kHz will have to be transmitted. Additionally it has to be considered that the user is limited concerning its own dynamics, or may even be further limited artificially. The DaVinci System (fig. 1.8 on page 15) as an unidirectional telemanipulator filters e.g. the high frequencies of the human movements to prevent a trembling of the surgical instruments.

12.2.2 Local Haptic Model in the Controller

A frequently used strategy being part of many haptic libraries is the usage of local haptic models. These models allow a much faster reaction on the user's input compared to the simulation of a complete object interaction (fig. 12.4). Such models are typically linearized functions dependent on one or more parameters. These parameters are actualized by the simulation at a lower frequency. For example each degree-of-freedom of the haptic system may be equipped with a model of spring, mass and damper, whose stiffness-, mass- and friction-coefficient is updated to the actual value at each simulation step, e.g. every $\approx \frac{1}{30}$ s. This approach does not permit the simulation of nonlinear effects in this simple form. The most frequent nonlinear effect when interacting with virtual worlds is the lift-off of a tool from a surface. Dependent on the delay of the actualization of the local model, the lift-off will be perceived as "sticking", as the tools is held to the simulated surface by the local model in one simulation step, whereas it is suddenly released within the next. Concepts, which model nonlinear stiffnesses, compensate this effect satisfactory. By making the additional calculations necessary for the local model, a significant data

reduction between simulation and haptic controllers is achieved. Distantly related concepts are used in automotive applications too, where CAN bus-systems are configured in their haptic characteristics by a host, and report selection events in return only.

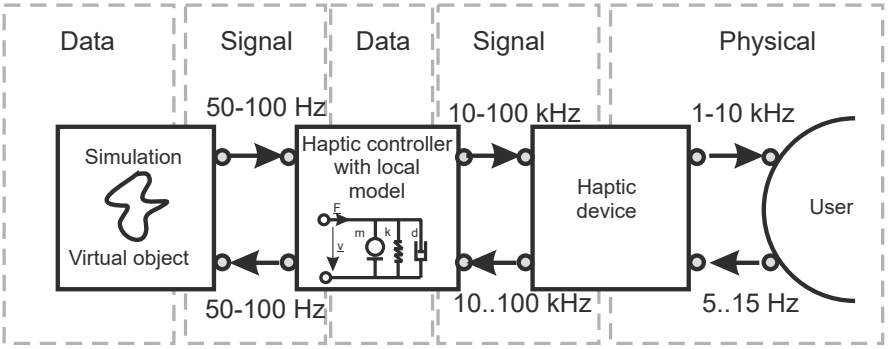


Fig. 12.4 Block diagram of a simulator with haptic feedback and a local haptic model inside the controller.

12.2.3 Event-Based Haptics

KUCHENBECKER presented in 2005 the concept of "Event-based haptics" [146] and brought it into perfection since. It is based on the idea to split low frequency interaction and high-frequency unidirectional presentation, especially of tactile information (fig. 12.5). These tactile events are stored in the controller and are activated by the simulation. They are combined with the low-frequency signal synthesized from the simulation, and are presented to the user as a sum. In an improved version, a monitoring of the coupling between haptic device and user is added, and the events' intensities are scaled accordingly. The design generates impressively realistic collisions with comparably soft haptic devices. As any other highly dynamic system it nevertheless requires a specialized driver electronics and actuator selection to achieve full performance.

12.2.4 Movement Extrapolation

Another very frequently used method for bandwidth reduction on the path to measure user reaction is given by extrapolation of the movement. Especially with simulators using local models it is often necessary to have some information about steps

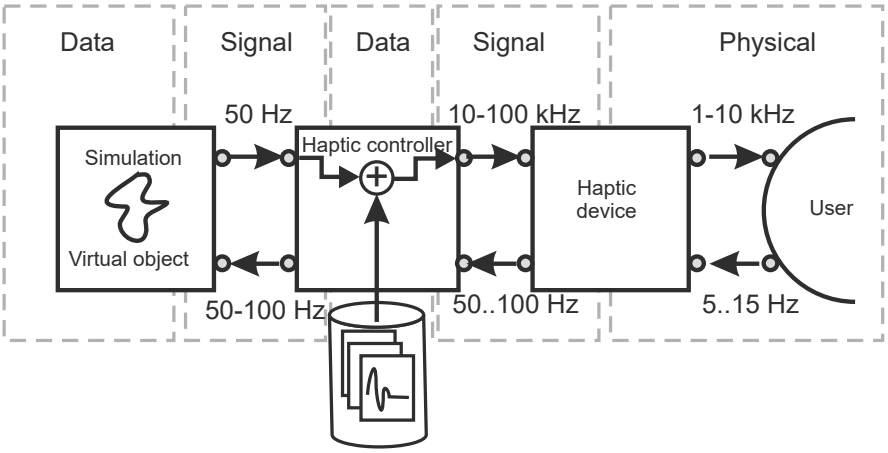


Fig. 12.5 Block diagram of a simulator with haptic feedback and with events of high dynamic being held inside the controlling structure.

in between two complete measurement sets, as the duration of a single simulation step varies strongly, and the available computing power has to be used most efficiently. The extrapolation becomes a prediction with increased latency and a further reduced transfer rate. Prediction is used for haptic interaction with extreme dead times.

12.2.5 Compensation of Extreme Dead Times

The working group of NIEMEYER from Telerobotics Lab. at the Stanford University works on the compensation of extreme dead-times of several seconds by prediction [178]. The dead-time affects both paths: the user's reaction and the information to the user, such as the haptic feedback generated. The underlying principle is an extension of the telemanipulation system, which is added with a controller of the manipulator and a powerful controller for the haptic feedback (fig. 12.6). Latter can be understood as an own simulator of the manipulated environment. During movement, a model of the environment is generated in parallel. If a collision happens in the real world, the collision is placed as a wall in the model, and its simulation provides a haptic feedback. Due to the time lag the collision does not happen at the position where it happened in reality. During the following simulation the collision point is relocated slowly within the model back to its correct position. By successive exploration of the environment a more detailed haptic model is generated. The method has the status of a research project.

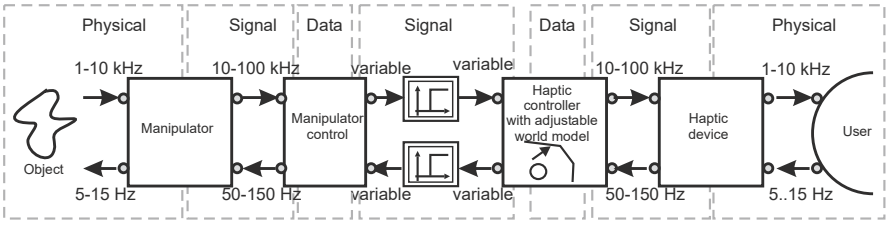


Fig. 12.6 Block diagram of a telemanipulator with compensation of long dead times by an adeptable world-model.

12.2.6 Compression

As any data data stream, haptic data can be compressed for reducing their bandwidth. This may happen based on numerical methods on each individual packet, however it may also be possible to make use of the special properties of the haptic human-machine interaction and haptic perception. The following list shall give a short overview about common approaches:

- A first approach for compressing haptic data is given in the situational adaption of digitalization on the path for measuring user reaction. SHAHABI [226] (University of Southern California, USA) compares different digitalization methods adapting their time- and amplitude-discretization on the actual movement velocity.
- Since several years now the working group around BUSS (TU-München, Germany) does intensive research on the perceptional impact of loss of resolution and bandwidth in haptic data streams [148]. They are coupling their research with the analysis of user-reactions, and base their algorithm on the psychophysical perception and an benefit-effort-analysis.
- The working group around EL SADDICK (University Ottawa, Canada) wants to achieve data reduction by standardizing the haptic interaction in a descriptive data format "HAML". It models the environment in a comparably little set of parameters, which gives advantages in tele-interaction applications with a larger group of participants. Especially varying data transmission paths can be compensated more easily on this abstract description, in comparison to classic telemanipulation approaches with a transmission of explicit forces and positions. As a by-product of this work, concepts for the unidirectional replay of haptic data in form of a "haptic player" are developed [47].
- Another obvious approach for compression is the usage of limitations given by haptic perception. The working group of KUBICA (University of Waterloo, Canada) demonstrates [296] an analysis of an interaction with a virtual environment at different velocities. The identified dependency of the force perception threshold on the velocity was successfully used as basis for data reduction.

- The working group of WERTHSCHÜTZKY (TU-Darmstadt, Germany) works on the measurement of haptic perception in the range of milli-Newton and micrometer resolution up to border frequencies of 1 kHz. They aim on building a database as a basis for compression with frequency dependent perception curves.

12.3 Technical Standard Interfaces

Most haptic devices are operated with personal computer or related systems. They offer a high flexibility in configuration as well as for research projects as for gaming- and design applications. Within this section different standard interfaces typical to PC hardware architecture are highlighted and discussed with respect to haptic devices applications.

Serial Port

The serial port is an interface, which is - dependent on the operating system used[1] - quite simple to be addressed. The serial interface of a home computer is based on the RS232 standard. This standard defines, beside several timing-aspects, the bits being encoded in between ± 3 V to ± 15 V for low- and high-levels. Compellable a connection to digital circuits with 5 V logic has to happen via a converter, such as the MAX232. Usually only two lines (RxD and TxD) are necessary for data transmission. The maximum specified data rate is 20.000 bit/s, whereas data rates far off specifications with 56 kbit/s or 128 kbit/s are possible too. Both, the data rate and the number of bytes carrying the actual data payload can be configured in a wide range. Two systems communicating with each other have to be adjusted in these parameters. A simple parity-control with one bit is also integrated. With respect to its data rate the serial port is absolutely suitable for the control of simple haptic devices. However due to its master-slave architecture a bidirectional data communication is connected with a large data-overhead as a results of the coordination of both units. As modern PC as not necessarily equipped with serial ports only, several USB-to-serial converters are offered at the market. They emulate a COM-port in software and transmit the data through the USB-connector and the typical 9-pole socket. The data rates achievable with these connectors are usually sufficient, but they show some unpredictable delays and time-loss due to software-emulation of the interface. This makes them hardly usable for time critical haptic applications.

[1] As a rule of thumb: with Linux or DOS-Systems a direct communication using the serial interface is a lot simpler than compared to Windows operating systems.

Parallel Port

The parallel port is, if it is still part of modern PCs, a 25-pole double row SUB-D-socket on the computer back plane. Similar to the serial interface its easiness of use is largely dependent on the operating system. In an idea case (with several Linux distributions and with DOS) the address of the port can be directly written with three consecutive bytes. In this case the first byte represents the eight data lines, the two following bytes are used for setting and reading of the control lines. The parallel port is set to work with 5 V logic levels with a maximum source current of ≈ 5 mA and sinking currents of maximum ≈ 20 mA. An overload current or leveling the pins to wrong voltages should be strictly avoided, as typical PCs do not show any protective circuitry. The data transmission of modern parallel ports is usually bidirectional, allowing read and write operations on the same eight data lines. But as the change between data transmission directions needs some time, the control lines are frequently used as input, whereas the data lines are operates as output. Writing to and reading from the port can be made with frequencies up to 100 kbit/s without much effort. An extension of the parallel port is given by the "Enhanced Parallel Port" (EPP) and the "Enhances Capabilities Port" (ECP). Whereas the ECP excels by Plug&Play-functionalities mainly, the EPP had been designed for an increase in bidirectional data transmission rates up to 2 Mbit/s. This increase was achieved by a Hardware-implementation of the data protocol. For a slave using the EPP - such as a haptic device - this additional hardware requires of course some more hardware on the device's end too, as the protocol has to be realized near to the interface. From the perspective of data rates, the parallel-port is highly suitable for haptic devices, especially in EPP-mode. Especially low latencies of $<100\,\mu$s between writing command and the availability of the data makes it very attractive. Only the dwindling availability of this port to standard PCs makes it necessary to use other type of interfaces.

By their flexibility in software drivers, the serial and the parallel port makes it possible to be interfaced as debugging ports to microcontroller circuits or to bus-systems such as I2C or CAN.

USB

The USB-port is a serial port with a predefined data transmission protocol. It contains two data lines, an electrical ground and a 5 V supply, which can be drained with up to 100 mA per device attached. According to the USB specifications this load can be increased up to 1 A, if the host accepts it. An extension of the standard named "Power USB" considers additional lines for higher currents and even larger voltages. The USB-clients receive an identifier when being connected to the bus, which marks the data packages, sent to or from them. Each USB-component has "Device Descriptor" uniquely identifying the manufacturer and the product. Additionally each device is classified in several standard classes. The "human-interface-device-class" (HID) is reserved for input systems. Devices with active, haptic feedback are

grouped in an own class of “physical-interface-devices” (PID). Each manufacturer of an USB driver circuit has to apply for a unique product id. Due to the requirements and the complexity connected to the implementation of an USB-conform protocol, it is recommended to use USB-interface circuits of the shelf for product designs with little quantities manufactured. Such interface circuits offer parallel or serial data lines to be interfaced by an own microcontroller, which itself does not have to bother about the USB interface any more. The USB interface can be operated in different modes. For the transmission of larger, time-crictical data volumes the “isochrone” transfer is most suitable. Its theoretical limit is given by the data packets, which are transmitted in micro-frames according to the USB 2.0 standard. The duration of a micro-frame is given by 125 μs. So called “full-speed-systems” are able to transmit up to 1023 bytes with each micro-frame. High-speed devices are able to transmit even three times more bytes per micro-frame. According to the specifications transmission rates of up to 40 MByte/s are possible with devices combining several isochronous endpoints in one unit. The data rate of isochronous transfer is optimized for unidirectional transmission only. In case of bidirectional communication the data rate is reduced accordingly. Nevertheless the speed of the USB-port covers any requirements given by USB devices. However two special aspects have to be checked in the context of the individual application:

- A micro-frame of 125 μs duration (8 kHz) is the upper limit of the available bandwidth. Without compression and decoding this gives the natural bandwidth limit according to NYQUIST of 4 kHz.
- The data rate has a tolerance of 0.1%.

FireWire - IEEE1394

FireWire, Apple’s brand name, according to the IEEE1394 standard is a serial transmission format similar to USB. In fact it is a lot older than the USB specification. The six-pole FireWire Connector includes a ground and a supply line too. The voltage is not controlled and may take any value between 8 to 33 V. FireWire 400 defines up to 48 W power to be transmitted. The data rates are - dependent on the port design - 100, 200, 400 or 1600 kbit/s. This is completely sufficient for any haptic application. Even fiber optics-transmission over 100 m distance with up to 3200 kbit/s are specified in the standard. The bus-hardware additionally includes a concept to share memory areas between host and client, enabling very latency less transmissions. Even networks without an explicit host can be established. The interface according to IEEE1394 is the preferred design for applications with high data transmission rates. Only the little propagation of this interface in personal computers hinders a wide application.

Ethernet

The capabilities of the Ethernet-interface available with any PC are enormous but largely dependent on the protocol used. Whereas the naked interface enables transmission rates of 10 Mbit or 100 Mbit or even Gigabit, the available data payload within the transmission is largely dependent on the interlacing of the underlying protocols. The very well known TCP/IP protocol has a header portion of 40 byte. The Ethernet protocol adds another 18 bytes for the Ethernet frame, and some more 8 byte for the whole packet, resulting in an overhead per packet of 66 byte. This packet may contain up to 1460 bytes of data. This is sufficient for typical haptic applications with respect to the available space per packet. Assuming a six-DOF kinematics with 16 bit (2 byte) resolution in their sensors and actuators, each packet has to carry only 12 bytes of data, with one packet for force-output and one for position-input. The number of bytes carried in each packet has a lower limit depending on the physical design of the network. In the area of home networks it is 50 byte, making it necessary to add arbitrary data on the example from above. A cycle of the haptic device example would transmit 232 byte, which is 1.856 kbit. With a 10 Mbit network and theoretical bandwidth of 8 kHz would be available. Even when considering that the data have to be extended with some additional overhead (address negotiations, status-information), this is still sufficient for many haptic applications. A disadvantage in using the Ethernet is given by the high efforts necessary for packet confection and protocol formulation, which would usually overload the computation power of standard microcontrollers. Additionally a high number of clients reduce the data rate within a network significantly. Using switches compensates this reduction to some extend.

Measurement Equipment and Multi-Functional Interface Cards

Measurement- and multi-function interface cards are a simple approach to interface to hardware designs. They are available for internal and external standard interfaces, such as PCI, AGB or PCMCIA. They are usually equipped with several standard sw-drivers optimized for their hardware capabilities. When considering a prototype design they should be considered in any case. Their biggest disadvantage is given by the data processing happening inside the hosting PC and within the restrictions of the operating system. Especially in combination with non-realtime operating systems like Windows the dynamics of controllers necessary for haptic applications may become not fast enough.

HIL-Systems

“Hardware In the Loop” (HIL) systems were first used in control engineering and compensate the disadvantages from multifunctional cards for rapid prototyping and interfaces to haptic systems. HILs include a powerful controller with proprietary or

open real-time operating system. The programs operated on these controllers have to be built on standard PCs and are transmitted as with any other microcontroller system too. Frequently the compilers allow programming with graphical programming language such as MatLab/Simlink or LabView. The processors of the HILs are connected via specialized bus-systems with variable peripheral components. Ranging from analogue and digital output over special bus- and actuator-interfaces a wide range of components is covered. HIL-systems are predestined for the always time-critical applications of haptics in design phase. But compared to other solutions they have a high price too.

12.4 Final Remarks about Interface Technology

The interface subordinates to the requirements of the system. Any realistic application and its required data rate can be covered with today's standard components. Only commercial- or company-interests may prevent the choice of a suitable interface for a haptic device. This is a complete difference to the situation at the beginning of the 21st century. At this time highly specialized interfaces were designed for haptic devices, to cover the high requirements on data transmission rates. Accordingly even today commercial products with own ISA or PCI interface cards can be found on the market. Other solutions require an EPP parallel port still. Nevertheless the design of controller circuit suitable for the USB protocol should not be underestimated. Especially its layout and programming for the still high data rates of haptic devices offers enough room for errors. Although the technical specifications are sufficient to fulfill the requirements, the first design and operation is far from being trivial.

"*This page left intentionally blank.*"

Chapter 13
Software Design

ALEXANDER RETTIG

The ultimate display would, of course, be a room within which the computer can control the existence of matter. A chair displayed in such a room would be good enough to sit in. Handcuffs displayed in such a room would be confining, and a bullet displayed in such a room would be fatal. With appropriate programming such a display could literally be the Wonderland into which Alice walked.
IVAN E. SUTHERLAND, 1965 [246]

A central application area for haptic systems is the so called "Virtual Reality". This names a concept of human-computer-interaction, which has developed rapidly during the last 20 to 25 years. The vision itself has been formulated already at the beginning of the development of computer graphics in the sixties of the last century by SUTHERLAND.

The focus of this chapter about software design is on this very application area of computer haptics in virtual worlds. Other application areas of haptic systems as the modification of haptic properties of manual control elements or telemanipulation respectively need a system structure, which is mainly defined by control engineering aspects as discussed in other chapters of this book (chapter 7 and 12).

Following a short overview to motivate the subject and provide some terminology and concepts different topics will be discussed from a software developer's point of view. According to the goal of this book the chapter at hand does not provide an exhaustive depiction but gives a basic understanding with creating interest for further activity with the topic in mind.

T.A. Kern (ed.), *Engineering Haptic Devices*, DOI 10.1007/978-3-540-88248-0_13,
© Springer-Verlag Berlin Heidelberg 2009

13.1 Overview About the Subject "Virtual Reality"

"Virtual reality"[1], or VR, terms a technology, which among others should fulfill the demand to recreate the natural environment as close to real as possible inside a computer simulation or to let unreal things become real and present them to the user quasi as if they were true. The displayed virtual environment should behave in the way the user would expect because of his experiences in the natural world. Ideally VR allows a perfectly intuitive handling with the computer which needs not to be learned.

Three main criterions have been identified during the research about virtual worlds which should be fulfilled as possible to achieve this goal: a quality of the display and form of presentation, which enables *immersion*, natural or intuitive *interaction* and realistic or at least plausible *behavior* of the displayed environment.

Immersion

Immersion, "diving" into the virtual world, happens when the user quasi forgets that he is interacting with a virtual world and the real world largely steps into the background. To achieve this it is crucial to activate as many senses as possible with the best quality as possible: mainly the dominant visual sense as well as the auditory sense, but also — as soon as manipulative interaction becomes relevant for the application — the haptic perception and finally the olfactory or even the sense of taste may be stimulated. The latter ones typically are only dealt with in a kind of exotic scenarios though. At least a variety of sense modalities should be incorporated, this is called *multimodal* presentation.

Natural Interaction

Experienced computer users are easily apt to consider operating the computer by a mouse to be natural and intuitive. However, actually so-called more or less abstract interaction metaphors find use which must be understood and trained first: is it "natural" to "grab" a virtual object by pressing a mouse button after an arrow symbol was moved onto the visual representation of the object on the screen by means of a movement spatially separate of it of a small plastic box (the mouse)? Really natural grabbing is certainly something else: the hand is moved to where the object is seen, although in a fraction of seconds contact is taken up with the sense of touch before finally the fingers are closed around the object. In addition, that the meaning associated to an interaction metaphor often is only partially obvious — one *opens* a document by dropping its icon onto the icon of a text processor program?

[1] There are concepts "Augmented Reality" (AR), "Mixed Reality" (MR) or even "Augmented Virtuality" as well, which name different kinds of mixture and embedding of real and virtual objects into a real or virtual environment. This differentiation is not needed in our context though.

Those doubtless are useful and reasonable concepts, but their semantics have to be learnt and understood explicitly. By using different input devices as data gloves up to full body tracking systems, which track the movement of all extremities, the VR research tries to provide actually natural interaction with complex data worlds.

Natural Object Behavior

The third criterion, plausible behavior, inter alia includes, that dropped things fall, liquids swash and flow, that you can not go through walls and objects in the virtual environment can not penetrate each other. These behavioral characteristics have to be provided by different simulations.

Especially the latter aspect of natural object behavior, the impenetrability of solid objects, is of importance: although it could be implemented using collision detection and more or less physically correct simulation in a purely graphical presentation, but then there is a possible discrepancy between the real movement of the user and the visual echo of this movement. It occurs, for example, if a virtual hammer stops on the surface of the virtual wall, while the real hand of the user has already deeply banged into the virtual brickwork. *Haptic rendering*[2] with force-reflecting robots makes it possible to overcome this shortcoming.

This reflection is not only of academic significance but motivated by a practical benefit for the user (or his employer). One benefit which by no means should be underestimated in this respect is the fun factor. Apparently this is directly commercially relevant in the computer games industry. In the wake of the rapid development of computer technology, which is very much driven by computer gaming, the personal computers available on the mainstream market by now are capable to display VR scenes in a complexity and quality, which was achievable solely with specialized graphics workstations from SGI, SUN and IBM still a few years ago. Thus VR technologies are about to enter the mass market: a lot of computer games in particular from the genre of so-called first-person shooters or motorsport simulations with their excellent graphics and realistic physics simulations can — albeit with some conceptual concessions — be called "Virtual Reality". Today the wealth of detail of these virtual worlds is enormous, immersion happens very quickly, you forget that you play, almost everything behaves plausibly. The interaction, however, is hardly natural or intuitive as long as the game is controlled by keyboard and mouse. Suitable input devices fill this gap partly: the motorsport simulation makes much more fun immediately when you have a vibrating steering wheel in your hands and throttles below your feet — and the driving performance gets better.

---

[2] The term *rendering* generally denotes the presentation process done by a computer system. Without further definition often the image generation process for the graphical representation is meant, but analogously the production of structured, information-carrying stimuli for other modalities is called acoustic, olfactory, or just haptic rendering.

This factor, better user performance, is also a driving force for the use of virtual reality in industrial or medical applications. Another major aspect is the possibility to save costs while increasing the quality of products at the same time by using virtual prototypes. The automotive industry as a pioneer in this area demands, for example, that vehicles could be presented ahead of the construction of any physical prototype in life-size in photo quality and accurate, verifiable and interactively adjustable lighting, enabling designers to decide on form details, the car paint and exact color of the interior. In the best case the software should be that intuitive, that nobody needs an explanation on how to use it.

In the same way virtual prototypes are used to examine, whether the vehicle can be assembled as planned. Since the worker himself is subject to this analysis also it suggests itself to let a real person do the virtual assembly. This approach is superior in many aspects to a non-interactive assembly simulation relying only on a human model.

To get applicable results from this kind of investigation, the virtual assembly scenario has to be as realistic as possible. Especially in situations with narrow construction spaces which are difficult to access, haptic feedback for collisions must be provided, because the mechanic needs and uses it more or less consciously for finding his way and accurately placing either the tool or the component.

Quite similarly haptic feedback vastly improves interactive simulations of medical interventions for the training of certain surgical procedures. Obviously the manipulative fine motor skills of the surgeon are of central importance during operations. Particularly in minimal invasive surgery much practice is needed for achieving these skills, since the physician can not interact in the operation field with his fingers, but has to do it in an entirely unfamiliar manner: the interaction is indirect using long thin instruments, which are inserted into the situs through small incisions. These punctures act as pivot points so that some of the movement directions are mirrored relatively to the movement of the hands. Additionally also the view is not direct, but via a screen showing an image taken by an endoscopic camera, which up to now still usually is monoscopic. The spatial orientation and hand eye coordination in this situation have to be learned and trained intensively. Traditionally models or animal cadavers are used, but more and more VR training simulators are establishing in this domain. Providing realistic visualization, simulation of soft tissue deformation, haptic feedback and tools for training analysis and evaluation they are much more flexible concerning different scenarios and much better suited for tracing the learning progress.

13.2 Design and Architecture of VR-Systems

13.2.1 Hardware Components

Generally the central hardware component of virtual reality systems consists of high-performance computers or clusters of computers which communicate via network. Connected to them are input devices and display systems using different interfaces.

The input devices mouse and keyboard as known from desktop computers are only of minor importance in VR applications. To fulfill the claim of natural interaction many other input apparatuses have been developed. Central are so called tracking systems, which offer absolute three-dimensional measuring of positions and mostly also of orientations in space, i.e. these provide 6 degrees of freedom (DoF). Using a tracking system it is possible to locate an interaction device exactly as well as to measure the (head) position and viewing direction of the user. The latter is amongst others necessary to achieve a perspectively correct visual rendering of the virtual environment or an appropriate spatial acoustics simulation. Several technologies are used for 6 DoF-tracking: magnetic tracking systems consist of sensors measuring the magnetic field, which is emitted by a field generator. Analogously ultrasound tracking systems have ultrasound emitters and sensors. Another class are camera based systems, where image processing methods are used to detect selected features of the real world (so called “marker”) in video streams. Their relative location is then calculated by taking the camera parameters into account. Usually such tracking systems operate with an update rate of 30 to 50 Hz. Next to the update rate also the latency between measuring and provision of the data in the VR system is important: if it gets too large, noticeable delays between user action and the visual result arise which destroy immersion: when for example the image follows with some time offset when the head is moved, immediately the impression gets lost, that the perceived visual stimuli come from objects with fixed location in space.

On the side of output devices several technical approaches are used for generating specific stimuli for the different sense modalities. For the auditory perception it’s the task of speaker systems, from headphones to surround sound systems and finally to speaker arrays by which precisely located sound events could be generated.

However, by far the most attention is given to the most dominant distance sense of man, the visual sense. Based on the physiological model of the trichromatic vision almost all optical display systems generate image frames at frequencies between 50 to 120 Hz, which are built of square picture elements (*pixels*) mixing the additive primary colors red, green and blue. The impression of continuity of motion is created by means of that the change of images is fast enough to trick the human vision system, which commonly happens from about 15 Hz [74].

To generate the impression of spatial depth stereoscopic displays are of exceeding importance in VR, because the presentation of a separate image with different perspective for each eye is the most effective way to stimulate the depth impression at least for people with normal seeing capabilities.

The sizes of graphical displays used in VR applications range from small optical units in data glasses or *head mounted displays* (HMD) to large tiled projections beamed by a bunch of projectors or projections onto the sides of a big cube which one enters into, a so called CAVE (recursive acronym for *Cave Automatic Virtual Environment*). Standard monitors or autostereoscopic monitors are also used, but they don't meet the demand for immersion very well.

Force reflecting haptic devices take an exceptional position amongst the hardware components of VR-systems: they are both input and output devices for they provide the input functionality of (mechanical) tracking systems in their *passive* degrees of freedom and also the (stimulus) output functionality by force transmission in their *active* degrees of freedom.

13.2.2 Device Integration and Device Abstraction

Many modern VR-systems implement the concept of *device servers* and *logical devices* for the integration of various input and output devices. A server is a computer or a running software process respectively, which offers a service to other programs. The service a device server provides is to supply input data on request or receive output data following a specified protocol[3]. The device server abstracts from the special characteristics of the physical device as far as reasonable. For example all 6 DoF tracking devices provide position and orientation data, but they differ from manufacturer to manufacturer with respect to the control sequences which are used to initialize and calibrate them and to trigger the data retrieval. These technical details could be hidden by the device server, which is adapted to the particular tracking system. It then as a service provides the logical device "position and orientation" comprising a standardized protocol for the transfer of control and payload data. This service furthermore can occur *network transparently*: the input or output device may be connected to another computer in the network, if the communication protocol between VR system and device server is designed accordingly and suitable configuration methods enable the VR-systems to gain access to the remote server.

The concept of logical devices disburdens the application developer who designs a virtual scenario from thinking about the exact type e.g. of the tracking system to be used. Later on it is even possible to replace an absolute by a relative 6 DoF input device like the so called SpaceMouse[4] without the need to change anything of the scenario, as long as the device driver of the SpaceMouse also implements the logical device "position and orientation". By the way this concept is quite common in the graphical "desktop" user interfaces of modern PCs: computer mice which are

[3] The term *protocol* refers to a scheme which describes the semantics of an otherwise abstract data stream and determines the order of its elements.

[4] The SpaceMouse consists of a knob roughly of the size of the palm of the hand. It includes sensors, which measure forces and torques the user exerts on it. The internal software converts these to relative movements.

connected via the USB port register themselves as *human interface device* (HID). Included device drivers take care, that the operating system can use the device as a "pointing device".

Quite similarly at least on a single computer the VR-system developer[5] needs not to worry about servers for the output of audio and image data as the operating system and standard libraries provide software interfaces to the respective device drivers, which control the hardware, i.e., the sound and graphics card. The latter on their own are responsible for sending appropriate signals to loudspeakers or monitors via their hardware interfaces. As soon as more complex output systems on distributed systems get involved though, the standard driver interfaces are not sufficient any more and VR-systems have to implement the needed functionality itself.

This holds e.g. for graphical display systems as tiled projection screens or CAVEs, which often are driven by clusters of computers, each one of it controlling one or more projectors. Server processes running on all of these computers collectively provide a service for the output of the composed image[6]. Similar concepts are used, when the acoustic output should be carried out by a separate audio workstation.

The device technologies in the area of VR are utterly varied, the development of concepts is still very dynamic and thus no common standards have been established yet. Every VR-system defines its own protocols for device connections[7]. It would be desirable, if there were cross-system classes for haptic devices similar to the USB-device class HID in the future. They should be based on popular interfaces which are capable of real time data transmission as e.g. IEEE 1394[8] (section 12.3).

Whereas small latencies in the data transmission via networks are mostly unproblematic for pure input devices, they become critical if they occur in the control loop of haptic devices. Therefore a network transparent device abstraction is not suitable for haptics without further ado. We go into that in more detail for the special case of the software renderer in section 13.2.5 and relating to the latencies in chapter 12.

Another aspect of device abstraction is of major relevance for haptic devices: the application developer is used to model the virtual environment in a Cartesian coordinate system and in the same way simulations are typically performed in Cartesian coordinates. The necessary conversion of the input data from device coordinates to Cartesian coordinates should be "hidden" from the application by the device driver or server as well as the transformation of the Cartesian output forces and moments into motor forces and moments and the appropriate control currents. Both conver-

---

[5] Very often one distinguishes between the system developer, who creates the software infrastructure (in our case the VR-system libraries and executables), and the application developer, who uses this infrastructure in order to model a virtual environment with object data and behavioral descriptions.

[6] On UNIX-like systems with graphical user interface based on the X-Windows system this concept is well established. Any computer with a running X-Server can display graphical output for all computers in the network which use the X protocol.

[7] There are some efforts to develop software libraries which define a standard server interface for various input devices, e.g. the free open source VRPN (*Virtual Reality Peripheral Network*) system.

[8] Apple's *FireWire*, SONYs *i.Link*

sions need internal knowledge about the kinematics of the device and the mechanical and electronic design, which almost without exception are irrelevant for the design of the virtual scenario.

13.2.3 Software Components

The central software instance of a VR-system is often called object manager or VR kernel. It manages the scene, i.e. all objects of the virtual environment, and controls the data flow between all hardware and software components.

In many cases the central data structure is a so called scene graph. It structures objects into groups in a transformation hierarchy and defines their relations. The whole virtual environment shares a global world coordinate system. Nodes which hold transformations define the relative location of their subordinate objects in the virtual world and parent child relations in the graph represent spatial and kinematics dependencies.

For example the four geometry objects, which represent the wheels of a vehicle, are connected via transform nodes to the node, which contains the car body, which itself is linked to the world coordinate system via a transformation node (fig. 13.1). By changing the parameters of the transform nodes over time it's easily possible to describe the dependent movements of the individual parts of the car, i.e. the rotation of the wheels around their axes relative to the car body as well as the superimposition of the forward movement of the whole vehicle.

The object manager furthermore manages the behavioral descriptions of the scene. These define the effects of an "event" on the objects and which subsequent events are initiated by it.

According to the implementation concepts of the particular VR-system these dependencies are coded directly in scripts or parameterizable modules or they are expressed in so-called behavior graphs or a combination of both is used.

After loading the scene description and setup of all data structures the VR kernel starts a cycle repeating upon system termination, which can be phased into 3 steps[9]:

1. Collecting all events
2. Propagation of the events through the behavior graph or triggering the event processing in the appropriate modules respectively
3. Display of the final state of the cycle on all output devices

The first phase amongst others comprises of requesting the data of the input device's via the device's servers depicted above. In the second phase the event data is passed to the various simulations and these calculate the next step. Since simulations may run concurrently as described below this phase ends with a synchroniza-

[9] Basically this process is the same for every interactive software system including all widely used programs that have graphical user interfaces — from word processors up to computer games.

tion leading to a consistent state, which is interpreted by the different renderers for the individual sense modalities to create the respective display data.

13.2.3.1 Events, Event Propagation and Behavior Graph

There are several sources for events: obviously user actions via input devices are events, which either occur at discrete points in time (finally you only click a mouse button now and then) or e.g. when using a tracking system at every time step, since tracking systems usually send data continuously.

Less obvious are those events, which originate from internal changes as the stepping forward of the simulation time, which triggers the calculation of a new state of animations or simulations. Moreover there are events, which result indirectly from user input or simulations as for example the collision of two moved objects.

On the implementation side events are represented by state changes of objects, i.e. by changes of the values of internal variables of them, which define their current properties.

The task of the object manager is to propagate event changes according to the scene description. This means, that it has to send messages to all concerned objects to inform them about relevant events, which they should react onto. To illustrate the underlying principles we will utilize the terminology of the ISO standard X3D [1]. As a rule even VR-systems which don't implement this standard work with similar concepts.

Besides the software object instances which correlate to those objects of the virtual environment, which are arranged in the scene graph and rendered by a renderer (visible objects by the graphic renderer, sound sources by the acoustic renderer, tangible objects by the haptic renderer), there are software object instances which exclusively provide behavioral functionality. Included amongst others these are simulations or animations, which for example store a predefined motion to be played back, controlled by a timer.

Both object types — scene and behavioral objects — are stored as nodes in the behavior graph. Therefore the scene graph could be considered as a subgraph embedded into the behavior graph. In addition to the parent-child-relations of the transformation hierarchy the behavior graph contains directed edges which model reaction chains of events. These edges are called *routes* (in X3D).

Routes deliver messages between nodes. In the simplest case a message is the transfer of a new value: an animation may for example generate a new position which should be transferred to the transform node belonging to the object to be animated. Within other concepts the message transfer correlates to the call of a function of the receiving node which may take several parameters. The state change initiated by the message in the receiving node very often results into the generation of new events, which again have to be transferred to linked nodes. This process of distributing or *propagating* messages is called *event cascade*.

We will give a simple scenario to exemplify this concept: for the replay of a time controlled animation a timer is instantiated in the scene. It generates a time signal

for each run of the VR application loop (the event is “a new loop has been started”). The event is transferred to an animation node by a message which contains the new time stamp. The animation node evaluates a stored function which maps the time value to a new position value (this evaluation is the subsequent event). The event is propagated to a transform node, which sets the new position value in the transformation it holds internally. When the next rendering is done (section 13.2.5) all objects which are descendants to the transform node in the hierarchy will appear at the new position.

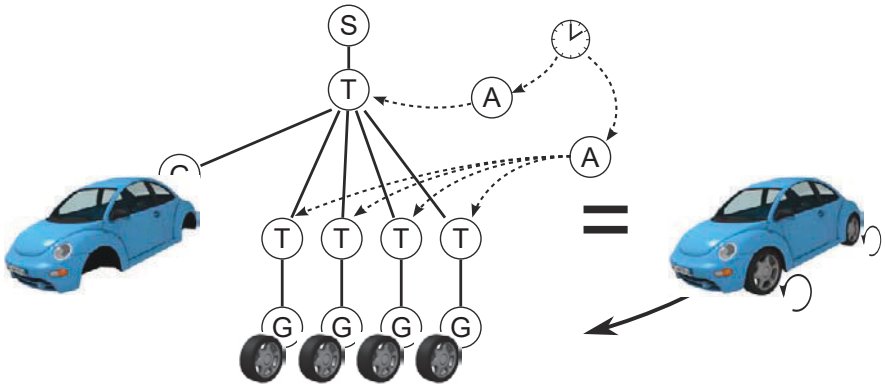


Fig. 13.1 Simple scene and behavior graph of a driving vehicle: S: root node of the scene, T: transformation node, G: geometry node, A: animation node, clock symbol: timer. Routes are dotted.

This composition of behavioral and scene objects make it possible to construct networks of dependencies which model very complex behavior inside a virtual world. Messages about events commonly are not only sent to one but many objects (by creating several outgoing routes). In the same way an object may receive many types of messages and react differently to them.

The control of this data flow is the responsibility of the object manager. In some constellations of the event cascade cycles could happen which have to be broken to avoid infinite loops, which otherwise would lead to a seemingly frozen system. The events which could not be processed for that reason in one time step but should not get lost are fed into the event cascade of the next time step.

13.2.4 Simulation

Despite of the diversity of behavior which could be created by the composition of simple behavioral elements it is not possible to describe everything what eventually is of interest in a virtual environment. Many desired behavioral properties of a

scenario cannot be defined by simple timer or event controlled processes but need more complex simulations. For example it is not feasible to predefine animations for every single activity during a surgical operation — taking hold of tissue, clamping, cutting at arbitrary spots, suturing — which would properly produce local deformations or even topological changes due to incisions. Rather this behavior of organs must be simulated in real–time and depending upon the use case different levels of physical correctness are necessary: if e.g. the focus is to train dexterity (motor skills) the reactions of the virtual tissue only need to be plausible and a kind of close to reality. If on the other hand a lesson is about improving diagnostic abilities (sensor skills and interpretation), where specific properties of anatomical structures should be examined for pathologic alterations, these have to be simulated with much higher accuracy and realism.

To implement the relations which have been identified during simulation modeling usually data structures and calculation methods are needed, which are not part of a general VR-system. Very often moreover there already exists an implementation of the physical model in a specialized library or system for physical simulation, which at the best could be reused and integrated into a given VR-system. To do so a concept is needed how the application loop of the VR-system exchanges data with the simulation component and synchronizes with it.

Simple simulations possibly may be implemented directly as nodes of the behavior graph to be processed within the application loop. The node encapsulates the interface of the simulation library, delivers the events it receives as input data to the simulation and feeds the results of it as subsequent events into the event cascade.

However this approach is only suitable if the simulation does not need too much computing power. Inside the application loop only a restricted amount of computation time can be allocated for performing a simulation step, since the various output systems need to be updated at regular intervals in order to guarantee a minimal graphical frame rate for example. Otherwise the virtual environment may feel sluggish or start to judder. Even if the simulation itself fulfills real time demands it may happen that the necessary computing resources are not available in the context of the complete system which also has to manage input and output devices and process the event cascade. In this case it is inevitable to decouple the simulation from the cycle of the application main loop.

To support the clarification of the problem we will at first discuss the term “real-time”, which is used more sloppy in the domain of computer graphics as e.g. in robotics: one essential temporal requirement of a graphical display is obviously to nearly always achieve update rates above a frequency of 15 Hz to create the visual impression of motion and avoid perceivable delays between user interaction and the reaction of the system. This often is already taken to be enough to refer to it as real-time computer graphics. More precisely though it would be to say that an application provides *interactive computer graphics* to emphasis, that the system is fast enough for fluid interaction.

When speaking of real-time in the simulation domain it is meant, that the simulated time or model time proceeds as fast as the real world time. Simulations can stretch or compress time (slow motion or fast motion) for to, for example, show pro-

cesses, which are too fast or too slow in reality to observe them. Simulations of this kind do not run in real-time. The underlying interpretation of "real-time" however does not define any constraint on the time needed to calculate one single simulation step.

On the other hand real-time capability of a control system especially in robotics denotes that the result of a calculation is available within a given period for use as control signal. If it is enough to fulfill this requirement on average and exceptions are largely harmless as long as they are not too frequent this is called "soft" real-time. Within "hard" real-time constraints the system must complete each calculation before a deadline in any case. The time to deadline may differ largely though: some systems only require a guaranteed result in a period on the time scale of seconds whereas others need to be clocked in microseconds.

To meet the requirement of hard real-time with multi-tasking (see below) control computers have to run under a real-time operating system like LynxOS, QNX or VxWorks. These take care, that processes with according privileges get the processor within a guaranteed time span for a defined period. Based upon this guarantee the processes running on the system can give real-time guarantees themselves.

The problem touched above — that not enough resources can be allocated for a simulation — now presents itself as follows: a simulation may, for example, not be able to perform a simulation step during the duration of one cycle of the application main loop, which may have to run with 15 Hz. Thus the simulation does not satisfy the soft real-time requirements of the VR-system. But it possibly meets the real-time condition for simulations if it for example can perform a step of 0.2 seconds of model time with a frequency of 5 Hz and a processor load of 50 %. Now the integration of the simulation into the VR-system becomes possible, if it is separated from the main loop into its own loop: about every three main application loop cycles the simulation can feed new data into the behavior graph via the node associated with it. Then the next simulation step can be started with new input data coming from other parts of the virtual world.

The concurrency of the VR main loop and the simulation is realized by outsourcing the simulation into its own so-called *thread*. Threads are a technology provided by the operating system, which allows it to let parts of a program run temporally independent from the rest of the program. This is similar to the concept of processes which is the basis of *multitasking*. Simply speaking the processor switches the program it processes in very short time slices. For the user it appears as if the programs would run at the same time. On modern processors with multiple processor cores or computer systems with multiple processors the execution of some of the programs (in case of processes) or program parts (in case of threads) actually takes place at the same time. For complex scenarios this concept is developed even further: it is also possible to outsource the simulation onto a separate computer or a computer cluster[10] as well. This is called a *distributed* system.

The next issue is the correct synchronization between the simulation and the rest of the system: each time data has to be exchanged between the two loops both

[10] *Cluster* in this context denotes a network of computers which collectively perform a task.

systems have to be in a consistent state. If a rotation matrix was only filled with new data to the half when it is transferred this would result into data corruption. Most likely the inconsistent copy of the rotation matrix wouldn't have certain essential mathematical properties. Problems of this kind are avoided by defining dedicated spots in the program where the involved threads wait for each other and thus only "approved" data is exchanged. We don't go into the problems of the synchronization of processes and threads in more detail here but as to be expected the effort to implement it becomes the more, the more complex the distribution of the system is: if threads run on *one* computer common memory in the RAM can be used for data exchange whereas in the case of cluster systems network communication including inherent latency problems gets involved.

13.2.5 Subsystems for Rendering

The main purpose of a VR-system is of course to display the simulated virtual environment. To do so various rendering subsystems have to interpret the current state of the scenario regularly and create an appropriate stimulus (a rendering) for every intended sense modality.

For the most obvious one, the visual or graphical presentation, this means, that about every 10 to 50 milliseconds the graphical rendering has to be triggered. Since it is only sensible to create a new image when a new state of the virtual scene is available usually the graphical rendering is directly integrated into the main application loop of the VR-system. After the event cascade has been processed and, if necessary, data from concurrent simulations has been synchronized into the main thread, the system starts a so-called traversal of the scene graph: the renderer recursively marches through the transformation hierarchy and thus by and by all nodes which contain information relevant for visualization are visited. The respective data is transferred to the basic graphics library[11]: affine transformation matrixes are accumulated from data for translation, rotation and scaling in transform nodes, material nodes provide color settings and textures, light nodes direction, color, brightness and attenuation of light sources and last but not least the triangles and other polygons stored in geometry nodes are "pumped"[12] to the graphics library. The library is responsible for transferring the data efficiently onto the graphics card via appropriate drivers. The graphics card finally performs the computation of a color for each pixel autonomously and creates the output signal for the connected monitor or projector.

It's worth to notice, that the graphics card controls the output device with a constant frequency (e.g. 60 Hz for TFT flat screens and preferably higher for CRT displays) which is completely asynchronous to and independent of the rate in which images are computed. For CRT displays a frequency of 85 Hz and above is recom-

[11] This basic graphics library usually is *DirectX* on Windows-Systems, on others the platform independent *OpenGL* (Open Graphics Library)

[12] It is spoken of "polygon pumps" in this context as well.

mendable for the reason that this avoids the perception of flicker by the receptors in the periphery of the retina, which are specialized in detection of movement: due to the functional principle of a cathode ray tube (CRT) the brightness of each pixel alternates, because the phosphor on the screen gets activated only for a short moment when the electron beam passes it on its zigzag way and luminesces decreasingly until activated again. On the one hand the afterglow duration should not be too long to enable the sharp display of fast movements without smear artifacts, but on the other hand a short afterglow duration abets the distracting perception of flicker. Flat screens or LCD or DLP projectors don't have this problem, because their backlight provides quasi constant brightness and pixel colors are created by filtering.

Hence for the graphical display one has to distinguish between image update rate and display refresh rate. The first one is determined by the cycle time of the VR application loop and has to be high enough, that single images fuse into motions, whereas the latter, higher one is determined by the graphics card and drivers and adjusted to the frequency, at which flicker gets imperceptible dependent on properties of the display hardware. Obviously there is no benefit of an image update rate of the VR-system above the refresh rate of the monitor or projection system and only would wastefully occupy computing power.

Some of the concepts of acoustic rendering are similar to those of graphic rendering. To take the location of a virtual sound source into account for sound synthesis the transformation hierarchy is evaluated and the orientation and position of the sound source as well as of the user determined. These parameters for the generation of stereo effects[13] are input into the audio library[14] together with *audio chunks* — the audio data fragmented into short pieces — from each sound source. The audio library forwards this data to the audio card, which generates the final output signal for the loudspeakers. The VR-system has to take care, that enough audio data is available for the audio card until the next update cycle. Otherwise distracting drop outs become audible.

Humans perceive frequencies above more than 10 kHz, to achieve a clear audio experience free of artifacts dynamics of about 20 kHz are preferable. Due to the enormous computing power it would need the generation of sounds by simulating the vibration behavior of the objects in the virtual scenario in realtime is impracticable. Most VR-systems therefore use the trick to combine the sounds from a set of prepared samples, even if that doesn't allow for the full diversity of sounds which is possible when interacting with real oscillatory objects. To compensate this the audio renderer may also interpolate between samples associated to different locations of the object the user interacts with [205].

Thus the requirements for the graphical and acoustic rendering are determined essentially by the biological properties of the senses. Neither the visual nor the

[13] For realistic surround sound simulation geometry information has to be included additionally.

[14] For example platform independent *OpenAL* (Open Audio Library) or specific for Microsoft *DirectSound3D.*

auditive sense retroact to the technical device, there is no coupling loop[15] which possibly could cause instabilities.

In contrast the haptic rendering has besides the *physiological* a *control theoretical* component too, because within haptic systems there is bilateral energy exchange between haptic device and user. Therefore as in general for controlling coupled systems hard real-time requirements have to be fulfilled when driving haptic devices, in particular force feedback devices: too large latencies in the control loop — i.e. control information is not updated for too long — destabilizes the system. Because typical haptic devices are tuned for low damping this could happen in a resonance frequency and carries the danger of self destruction of the system or hurting the user. This leads to the postulation, that high control frequencies of above 1000 Hz (chapter 12) should be achieved for the most frequently used devices with low impedance (chapter 5) or when high force gradients are wanted. Besides this technical reason the high frequency has implications for the quality and fidelity of the haptic impression. Since the haptic sense perceives vibrations of above one kilohertz (section 4.3) people find haptic systems which are able to display information in this range to be more realistic. This especially applies for the extreme cases, free space movement and stiff contacts.

13.2.6 Decoupling of the Haptic Renderer from other Sense Modalities

The visual presentation needs update rates, which are oriented to the temporal resolution of the visual sense. Even for the perception in the peripheral field of view and for extra fast movements 70 Hz is quite enough. For professional applications even lower frequencies of about 20 Hz suffice, because there virtual objects as well as the users usually move much slower than in computer games. Thus the requirements to a graphical renderer with respect to the update rate are about the factor 10 to 100 lower than to a haptic renderer. It follows, that the haptic renderer has to be decou-

[15] Strictly speaking even in this case there is a feedback loop which is closed via the interaction of the user with the virtual world. The instabilities, which occur in practice, usually don't result from specific properties of the graphical or acoustical rendering, but from unstable physics simulations. Though resonances occur *in* the presentation, but not *because of* the representation. An object may jitter on the monitor, but the luminous flux from the monitor doesn't get chaotic thereby or overdrives and endangers the user either.

This may be different with acoustical rendering. There it indeed could happen, that the user's interaction and a slow synthesis of the sounds lead to resonance effects, which — if the sound system is accordingly powerful — may cause acoustic overdrive. But this problem is mostly uncritical due to the limited capacity of the sound hardware.

pled from the cycle of the main application loop of the VR-system and the other, slower renderers, for which such high update rates are neither feasible nor fruitful[16].

The decoupling of the haptic rendering loop is implemented using the same principles as the concurrency of simulations (see sections 12.1 and 12.2): following the initialization phase of the VR application the system starts a thread, which concurrently to the application main loop executes the haptic rendering loop with high priority and frequency. It consists of the query of position data from the haptic device, collision detection, contact classification with force simulation and finally actuating the device driver with the calculated forces and moments. At the beginning of its cycle the main thread reads position data of the haptic device from the haptics thread to feed it into the event cascade. The data flow in the opposite direction is more intricate:

In complex scenes the collision detection is the time critical part of haptic rendering. Therefore it is an efficient approach to reduce the amount of data on which collision detection has to be performed in the haptics thread: if the main thread only supplies the haptics thread with a local extract of the scene, the *global* collision detection could run with the lower frequency of the main loop whereas the high-frequency collision detection only "sees" object parts, which currently are close to the haptic probe. After processing the event cascade the main thread estimates an area from the current position and movement of the haptic probe[17], which includes the haptic probe and the predicted path of it at least up to the next collision cycle of the main loop. The estimated area should have the smallest volume possible, but it is better to overestimate a little than to underestimate. All polygons of the scene which intersect the area are collected and pumped to the haptics thread and rendered by it. Collecting the relevant polygons itself can be implemented using collision detection algorithms by approximating the area by a polygonal hull object, which then is tested against the polygons. Obviously the collision detection in the main thread has to be done regularly and in not too long intervals, because otherwise the prediction of the motion of the haptic probe gets too inaccurate and the portion of the scene which has to be provided to the haptics thread gets too large. This again would make the local collision detection in the haptics thread more costly.

The data flow between main application thread and haptics thread has to be designed carefully, thus that the haptics thread never has to be stopped longer to wait for new data from the main thread. During the data exchange between threads it is inevitable to protect the memory areas which currently are written by one thread from any, even reading access by other threads, because the data is not consistent during the write operation. To achieve this the memory area is locked by the writing thread using operation system functionality. Any other thread running into a statement during the program execution which needs access to the said memory area, has

[16] Although the limit frequencies for acoustic rendering are above those of haptic rendering it could stay in the main application loop, when the sample-based approach is used. Only a physically based sound simulation would have to be decoupled, but this should not be discussed further here.

[17] The velocity of the haptic probe can be calculated from the position and duration of the last cycle in the main thread of course, but it's more accurate, if this is done in the haptics thread also.

to be stopped until the writing thread unlocks it again. To prevent that the haptics thread gets thwarted by the main thread when it writes the local scene portion the locking phase has to be kept short.

A good approach for this is to use *double buffering*: while the haptics thread is rendering a local copy of the scene in one memory area, the main thread fills a second memory buffer with the next local scene copy. Not before this copy is transferred completely to the buffer, the haptics thread is notified, that it should start using the data in the second buffer for rendering and release the first one to be refilled by the main thread. Only for the transfer of the notification and the buffer switch both threads must be synchronized. This should be done that way, that the main thread has to wait for the haptics thread if necessary but not vice versa — on the one hand because of the higher criticality of the haptic rendering, on the other hand because the haptic loop reaches the synchronization point in the program more often due to the higher loop frequency anyways and therefore the average waiting period of the main thread is much shorter than it would be for the haptics thread conversely.

The concept of the local scene copy is also used with success, if the haptic rendering loop runs on another computer than the main application loop. For example there are haptic devices like the *FCS HapticMASTER* (fig. 5.9 on page 102), which are sold with a separate driver computer on which the haptic loop is executed. In this case it is not necessary any more to have a dedicated high-frequency haptics thread in the VR-system itself. Apart from the differences in the inter process communication on operation system level the same principles are used for data exchange and synchronization.

13.2.7 Haptic Interaction Metaphors

As aforementioned shortly, the term *interaction metaphor* calls a concept in the context of "human-computer interaction" (HCI), which defines how a user action is linked to the reaction of the computer system. The way people work with a computer today by moving the mouse to control a little arrow and to click on mouse buttons to trigger actions is based on a complex interaction metaphor, which consists of many smaller interaction metaphors. One of them is e.g. the metaphor "drag and drop": pointing to an object, click, hold the button moving the mouse and thereby dragging the object onto another one as the icon of the tray, let it go. This interaction metaphor is patterned after the natural grasping, moving and releasing of an item.

For the hapticly enriched interaction with a virtual environment on a very basic level two interaction metaphors can be discerned.

Ideally if a perfect haptic device was available, the user could directly interact with all objects of the virtual environment including the sense of touch: when grasping in the very moment of the contact between fingers and hand to the virtual object a haptic perception would occur. The fingers of the hand as well as any other part of the body would be totally free in movement without resistances except for those,

which are defined by the virtual world. Everywhere at the body haptic stimuli could be provided. This shall be called *direct* haptic interaction.

Partially it is possible: the end-effector of the first versions of the haptic device PHANTOM for example is a small plastics cone like a thimble where the user puts his fingertip into to move the TCP. The force feedback of the system leads to the impression that you (almost) directly touch the surface of virtual objects.

To extend this to the whole hand it would be necessary to have many effectors, one for each finger or even better one for each phalanx, several for the palm and the back of the hand as well. Some developments of exosceletons lead into this direction, but there are no devices yet which could generate an even only roughly realistic impression. The software side seems to be the less demanding part of this challenge: the hand can be modeled in sufficient granularity using todays software technologies. The collision detection and force simulation between hand model and virtual environment is an ambitious, but not an fundamentally impossible task. On the side of the hardware development even the exact tracking of all possible hand movements is extremely difficult: each finger can be moved in 4 degrees of freedom, the thumb in 3 and additionally multiple motions of the metacarpus are possible. Conversely for each of these degrees of freedom actuators of best quality have to be plugged together in a very confined space. Even if only the fingertips are considered the technical complexity is enormous.

Because of that another interaction metaphor comes in handy which is used with all widespread haptic devices: a kind of indirect interaction with the virtual environment called *tool-handle metaphor*. Within this concept one does not directly touch the virtual environment with the fingers, but indirectly via the item one holds in the hand. It is as one would touch the environment with the tip of a pen instead of the fingertip.

This admittedly is a severe restriction but it still allows utterly valuable haptic interaction while requiring much less effort with regard to the device technology. First of all in many applications in reality the haptic perception is indirect also. That's obvious for minimal invasive surgery but also eating with a knife and fork, feeling out the notch of a screw with the screwdriver, writing or drawing with a pen or brush, modeling clay with tools and doing woodwork with a chisel are examples. When simulating these applications in virtual reality it is feasible to equate the end-effector with the tool. If the end-effector of the haptic device could be exchanged then it is possible the effector can be a clone of the grip of the real tool. It's noticeable, that the applications mentioned above are of less explorative nature but are more related to the modification of objects.

But the tool-handle metaphor is also useful in different problems: whenever it's suitable to equate the end-effector with a (substantially stiff) virtual object which should be moved around in the virtual environment it's a good choice to use a real representative of it as end-effector. In combination with rapid prototyping this approach is of special interest for assembly simulation as aforementioned in the introduction. The part which should be tested for assemblability in VR can be produced by stereolithography and mounted as end-effector. The mechanic simply takes and moves it as in the real assembly situation.

The choice of the interaction metaphor plays an essential role for the design of the device as well as for the planning of the software application and the algorithms to be used.

13.3 Algorithms

Depending on the purpose of a visual presentation of virtual objects, different properties can be used, others can be ignored. Many aspects like contour, color, reflective properties, transparency, local color-variations (texture) contribute to a realistic visual impression. The most important properties however are those, which allow to recognize the shape of the object. Consequently first efforts were aimed at contours and shape-defining edges during the historical development of computer graphics. With increasing computing power of the hardware and improved capabilities of the software the generation of brightness variations, color and other visual object properties could be realized.

A very similar approach can be found with haptic rendering, enabling the user to perceive shapes first and material properties second. Such impressions are generated when body parts of the user — usually finger or hand — hit a mechanical resistance during explorative movements. This resistance has to show certain properties: with regards to haptic rendering a spatial and temporal coherence has to exist, it has to depend on the exerted force, and there must be a spatial relation between user and the resistance. Consequently a haptic device has to be able to generate varying forces at different positions in space.

To get a three-dimensional representation of the involved objects they must be described in a mathematical model first. In general the same representations can be used as in 3D computer graphics applications. Instead of those attributes necessary for a visual representation, properties like stiffness, elasticity (hardness), roughness or stickiness stand in.

The dominating representation to define spatial structures within the so called “real-time”-computer-graphics are sets of triangles approximating the surface of an object. Even CAD-systems providing surface representations of higher order[18] for modeling generate a triangle-based intermediate representation to render on the screen. This is a concession to the fact, that today’s graphic hardware is highly optimized for the handling of triangles. In special application areas (e.g. image acquisition technology in medical context like computer- or magnet-resonance-tomography) direct volume models are visualized via *Direct Volume Rendering* (DVR), making use of regular 3D-grids of scalar- or vector-data (voxels). Both representations may be used for haptic simulations too, however additional data structures for collision detection (section 13.4) have to be generated.

Beside the virtual objects, the interacting user has to be represented within the virtual environment too. In the graphical rendering the corresponding counterpart of

[18] e.g. non-uniform-rational-b-spline tensor surfaces, NURBS

the eye would be the camera, characterized by position, orientation, aperture angle and other parameters. For the haptic rendering at least the hand — for the most frequent type of interaction in virtual worlds, which is manual interaction — is represented by a point in space. Its coordinates are derived from the position of the handle of the haptic device, the *Tool Center Point* (TCP). However this point in the virtual environment has to be distinguished from the TCP in the real world. In general coordinate transformations are necessary to match the position data from the device's TCP to the position of the point in the virtual environment. A simple example would be a molecular construction kit with haptic feedback, making it necessary to scale movements by several orders of magnitude. The representation of the TCP in the virtual scene is called *haptic-interaction-point*, or short HIP, in the following.

For some haptic rendering algorithms another object beside the HIP exists to represent the user-interaction. This object may be point-like or more complex. It is called *haptic probe*. Whereas the HIP always represents the position of the TCP within the virtual scene, the haptic probe is the object used to figure out the contact situation. In so called "penalty"-based approaches (section 13.3.2) the position of HIP and haptic probe are identical, whereas they may differ in other methods.

The most basic component of a haptic-rendering algorithm is a collision detection tuned to the object representation. It calculates the localizaton of the contacts between user and virtual object. The general concepts of collision detection are subject to section 13.4.

The collision detection is strongly integrated within the total structure of haptic-rendering algorithm. Put in a scheme, the cycle of haptic rendering can be systemized as follows:

- Position-/orientation detection of the user's interaction
- Collision-detection and evaluation
- Simulation
- Force or position output via the haptic interaction device

The following sections present several methods to render virtual objects haptically.

13.3.1 Virtual Wall

The simplest haptic model enables the virtual touch of a smooth plain surface with a dot-like probe, as it would happen with e.g. a hard pencil's tip. For this purpose a haptic device is actuated that way, that no resistance is excerted as long as the user moves in free space. For simplification we assume the free space to be the half-space with positive x-coordinate. Then any movement of the HIP through the yz-plane would get immediately slowed down by a counter force **F**. The user perceives this constraint as contact with a virtual wall.

The calculation of the counter force in the simplest format is:

$$\mathbf{F} = \begin{cases} K \cdot (x, 0, 0) & \text{for } x < 0 \\ (0, 0, 0) & \text{else} \end{cases} \tag{13.1}$$

with K being the proportional factor which controls the stiffness of the wall. Conceptional a spring according to HOOK's law is simulated, being effective in the normal direction of the yz-plane only (fig. 13.2). Movement components are not constrained in parallel to the plane. This makes the plane feel perfectly smooth. A large spring constant K results in a fast force increase, making the virtual wall feel stiff, whereas a small K generates the impression of a soft wall, feeling like made of foam plastic.

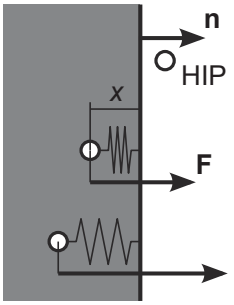


Fig. 13.2 Virtual Wall

This model of the virtual wall also could be interpreted as force field. In front of the wall (the yz-plane) the force-field is zero. Behind the wall surface the potential of the field increases proportionally to the distance to the surface.

This model can be easily generalized to a plane of arbitrary orientation and position. If a plane is defined in the HESSE's normal form with a normalized normal-vector $\mathbf{n}$ and a distance d of the plane to the origin according to

$$d = \mathbf{xn} = \begin{pmatrix} x_0 \\ x_1 \\ x_2 \end{pmatrix} \begin{pmatrix} n_0 \\ n_1 \\ n_2 \end{pmatrix} \tag{13.2}$$

the calculation of the force $\mathbf{F}$ for a position $\mathbf{p}$ of the HIP is given as:

$$\mathbf{F} = \begin{cases} K(d - \mathbf{pn})\mathbf{n} & \text{for } \mathbf{pn} < d \\ 0 & \text{else} \end{cases} \tag{13.3}$$

The formulation of the control loop of the haptic rendering for the virtual wall is given as:

```
repeat:
  read HIP position p
```

```
s = d - DotProduct(p, n)
if s > 0:
  F = ScalarMultiplication(K * s, n)
else:
  F = 0
actuate HIP with F
```

If K increases, the control loop has to be processed with an increased frequency to guarantee stability of the whole system. As a result of dead-times from discretization and other system components, the system tends to generate more energy with increasing product of K and δt. This makes the virtual wall to feel not as passive anymore as it is "expected" from a real wall. If the damping is not sufficient, the resulting energy makes the system oscillate in its resonance which can be perceived easily (section 7.3.3).

If an additional damping hardware or some closed-loop control in an external unit is not possible or reasonable (section 7), some additional damping in the control-loop of the haptic rendering may increase the stability of the overall system. One option is to add some damping force to the spring-force, actuating against the movement of the HIP and being proportional to its velocity $\mathbf{v}$ with a proportional factor D, the damping constant. Consequently a spring-damper system

$$\mathbf{F} = \begin{cases} K\mathbf{q} - D\mathbf{v}, & \text{for } s > 0 \\ 0 & \text{else} \end{cases} \tag{13.4}$$

with $s = d - \mathbf{pn}$ and $\mathbf{q} = s \cdot \mathbf{n}$ results. The vector $\mathbf{q}$ describes the distance between the HIP $\mathbf{p}$ from the plane in the reverse normal direction, being the penetration depth in the virtual wall. To still render an ideally smooth surface it has to be made sure to use the normal component of the velocity for the damping only: $\mathbf{v} = \dot{\mathbf{q}}$. If otherwise $\mathbf{v} = \dot{\mathbf{p}}$ is used, the impression of friction on the surface is generated. This impression however is not realistic, as it is independent of the contact force and shows no transition from static to dynamic friction.

In the implementation[19] of the control loop the velocity $\mathbf{v}$ is approximated by the differential quotient $\frac{\mathbf{q}_i - \mathbf{q}_{i-1}}{t_i - t_{i-1}}$ of the values from the actual time-step (i) and the time-step before ($i-1$):

```
initialize q_old, t_old
repeat:
  read HIP position p
```

[19] In the implementation of time critical systems such as haptic control loops care should be taken not to make repetitive calculations of similar or unnecessary values. Modern compilers are able to optimize such calculations out of the machine code. Some manual "optimizations" may even hinder the compiler to generate an optimized code. This makes it necessary to optimize the algorithm first by some complexity analysis. Afterwards it is implemented and only after careful considerations detail-optimizations should be made to the code. Anyway tests should be made to prove that the optimized code is performing better than the non-optimized version. Only very experienced programmers are able to compete against good compilers in this discipline.

```
read current time t
s = d - DotProduct(p, n)
q = ScalarMultiplication(s, n)
if s > 0:
  v = ScalarMultiplication(1/(t_old - t),(q_old - q))
  F = ScalarMultiplication(K, q) -
      ScalarMultiplication(D, v)
else:
  F = 0
display F at HIP
q_old = q
t_old = t
```

It is obvious that even with damping included the duration of a control cycle should not get too long, as the quality of the velocity approximation degrades. Nevertheless a clever tuning of the parameters K and D may achieve some good haptic quality even with slow sampling rates. If oneself is in the lucky situation to cooperate with the hardware-engineers and to be able to influence the hardware architecture, it is recommended to examine the whole signal processing chain including hardware and control architecture according to chapters 5, 7 and 12. Usually this makes it possible to find a much more optimized solution than tuning the rendering algorithm only.

13.3.2 "Penalty" Methods

The haptic-rendering method of a virtual wall in the previous section follows a so called "penalty" based approach. These methods are characterized by a cycle starting with a collision-detection to check, whether the HIP is within the boundaries of the object — whether it is "in contact"[20] Afterwards the contact is classified, and a repulsion- or "penalty"-force is calculated. This force is directed against the impact direction and tries to push the HIP out of the boundaries of the object.

The "classification of contact" characterizes the penetration depth and direction as a basis for the calculation of the penalty-force. This step is trivial for the virtual wall and resembles the calculation of the vector $\mathbf{q}$. The function of the collision detection in this simple model is covered by the query whether $s > 0$. Collision detection and classification with complex objects requires much more efforts. In fact the most time-critical part of real applications is given by the implementation of the collision detection.

As written before, penalty based methods with dot-like haptic probes do not have to distinguish between their own position and the position of the HIP. An obvious definition of the penetration depth and direction is given by the minimum distance

[20] The exceptional case of the HIP being exactly on the virtual surface at the moment of measure is very unlikey due to the time discretization. Nevertheless formally it is part of the contact-situation.

of the HIP to the surface of the virtual object (fig. 13.3). The repulsive force shall try to move the HIP to the nearest surface point. This allows the calculation of the penetration depth and direction for every point within the object in advance, as it does not change during interaction. In this case the classification of the contact can be done very efficient in a preprocess, which will be described for the case of the polygonal representation of virtual objects as polygonal surfaces in the following.

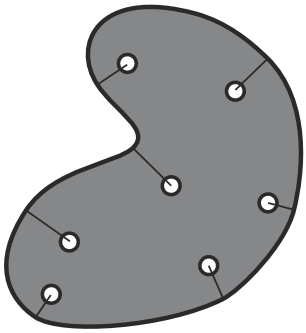


Fig. 13.3 Minimum distance between several points and object surface

Beside some negligible exceptions, every point of the object's interior has a minimum distance to exactly one point of the object's surface. Polygonal surface models have the features vertices, edges, and surfaces. All points of the object's interior which have an equal distance to the same feature — the same edge, surface or vertex - can be summarized in a common region, called VORONOI *region* [154] of the attribute. The VORONOI regions of all features partition the interior of the object (fig. 13.4). The partition is computed via so called *medial-axes-transformation* (MAT), which will not be detailed further in the context of the book. The term origins from the name of the boundary surfaces of the said regions which are called *medial surfaces* (and in the two-dimensional case *medial axes*). They consist of the set of points which have minimum distance to more than one single surface point.

Each Voroni-region of a feature of the object's surface is bounded by the feature and medial surfaces. Convex polyhedrons have flat medial surfaces, whereas in general VORONOI regions of concave edges or vertices show curved medial surfaces. To represent VORONOI regions they usually are approximated by polyhedrons themselves.

During runtime of the haptic rendering algorithm the VORONOI region has to be identified which contains the HIP. If the HIP is exactly on a medial-surface during one cycle (a usually purely theoretical situation), the above definition of the penetration direction is not unique. In this case one neighboring area has to be chosen, e.g. that one lying into the direction of the actual movement. The search for the VORONOI region can be done with a classic hierarchical bounding volume algorithm (section 13.4).

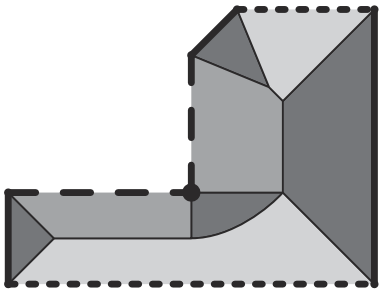


Fig. 13.4 Inner VORONOI regions of an object. Areas in light grey belong to dotted features, medium grey to dashed features and dark grey to drawn through features (and the thick dot representing a convex vertex).

Once the VORONOI region has been identified, the rest of the classification algorithm is simple: the vector from the HIP to the closest point of the adjacent feature is calculated, which is a standard procedure in any geometry calculation.

Penalty-based approaches have some fundamental difficulties though. The classification of the contact and the identification of the penetration depth and direction consider the actual situation only. This results in instabilities at the borders of the VORONOI regions with perceivable and frequently annoying changes in the force-direction and intensity. Especially with thin structures and relatively small spring-constants this effect becomes very dominant, when the HIP switches into the region corresponding to an object surface opposite to the original penetration point. As a result an unexpected change in force direction happens; the original resistance is replaced by a force acting in the direction of movement pushing the HIP out of the object on the opposite surface. Figure 13.5 visualizes this situation.

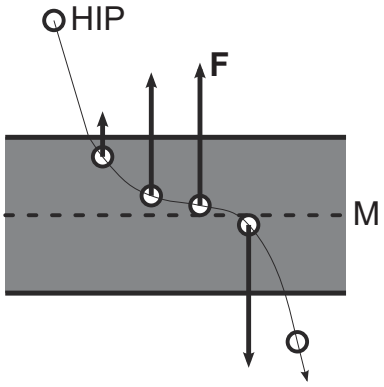


Fig. 13.5 "Tunneling through" thin structures: sudden inversion of the force direction when crossing the medial axis (M)

This conceptional disadvantage — not to cover the "history" of the contact — and its consequence leads to the constraint-based methods presented in the following.

13.3.3 Constraint-based Methods

The principle of constraint-based methods is based on the approach to model the virtual objects of the scene as borders of the free space impenetrable for the haptic probe. Whereas the user-guided HIP is able to penetrate an object, the haptic probe is kept on the surface of the object and additionally constrained to keep the minimal distance to the HIP. From the relative position between HIP and haptic probe a force is calculated pulling the HIP in direction to the haptic probe. This force is made perceivable for the user when displayed by the haptic device.

In 1995 ZILLES and SALISBURY [299] proposed one representative of this algorithmic class called *god-object*[21] algorithm (also named *surface-contact-point* algorithm). This algorithm is one of the most central algorithms for haptic rendering and is implemented in many haptic-rendering libraries. It solves the major problems of the penalty-based methods described before, and enables a one-point-interaction with polygonal surface models. The algorithm is suitable for the control of haptic devices with three active translatory degrees of freedom.

The method can be described best when following the trajectory of the HIP and a virtual probe in contact with a virtual object (fig. 13.6). Initially the HIP is in free space. The location of the haptic probe $\mathbf{p}_0$ is identical to the HIP. Now the user moves the handle of the haptic device. The algorithm checks whether the haptic probe is able to follow the HIP by performing a ray intersection test against the scene. If the movement does not cross an object's surface, the haptic probe is set to the new HIP position and the next cycle begins. Otherwise the crossed polygon's supporting plane[22] E_1 is set as a constraint to the haptic probe and limits its movement: the point $\mathbf{p}_1$ on the plane has to be identified, which has the minimum distance to the HIP. This is the new candidate for the position of the haptic probe. But before setting this point it has to be checked whether the passage-way between $\mathbf{p}_0$ and $\mathbf{p}_1$ is not blocked. With non convex objects it may happen that a second polygon has to be considered, resulting in its supporting plane E_2 becoming another constraint to the position of the haptic probe. In a three-dimensional case the position of the probe is limited to the intersection line of both supporting planes. The new candidate $\mathbf{p}_2$ for the position of the haptic probe is the point nearest to the HIP on this line. A third time the passage-way between the points $\mathbf{p}_0$ and $\mathbf{p}_2$ has to be checked. It may be that the supporting plane of another polygon limits the possible positions of the

[21] The haptic probe is called "god-object" by the authors, another typical term is *virtual proxy*.

[22] The *supporting plane* is the infinitely extended plane the points of a flat polygon lie on. If the vertices of a polygon are not in one plane, it has to be subdivided into flat polygons before, usually into triangles. This step should go into a preprocess and will not be discussed in further detail in this book.

haptic probe on the collective intersection point of all planes. Otherwise $\mathbf{p}_2$ is the new location of the haptic probe. As three supporting planes constrain the possible position of the haptic probe to one point already, another test in the same cycle is not necessary.

After the new position of the haptic probe was found, a spring-damper system (equ. 13.4) between HIP and haptic probe is used to calculate the force to be displayed haptically.

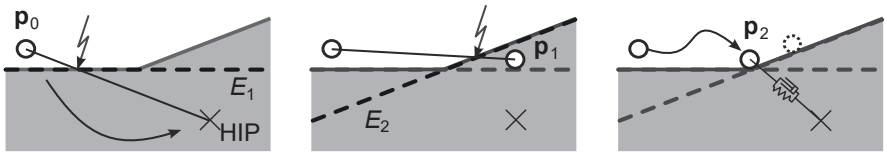


Fig. 13.6 Procedure of the god-object algorithm (diagram for 2D case). The optimum position of the haptic probe is dotted.

It should be noted that the algorithm does not guarantee, that the haptic probe reaches a local minimum distance to the HIP within one time step. Such a situation is displayed in figure 13.6, where the optimum point lies on the second supporting plane and not on the first one. Nevertheless the algorithm limits the haptic probe to the intersection lines of both supporting planes. This is not critical, as already within the next time-step the haptic probe moves into the direction of the optimum point. Practically due to the high update rates no artifacts can be perceived, as the method converges in a few time-steps to the optimum position of the haptic probe. The theoretical disadvantage that the algorithm in some rare cases may need several time-steps to converge is compensated by the advantage that it terminates for one time-step after three tests at the latest: a result is guranteed to be computed within a reliable time-frame.

In this algorithm the largest part of the runtime is usually consumed by the ray intersection tests for the identification of constraint planes. This makes it necessary to use sophisticated collision detection methods. On the contrary the minimization of the distance while satisfying the plane-constraints is less demanding. It is done as follows: the distance between haptic probe $\mathbf{p}$ and HIP $\mathbf{q}$ is minimized by searching the minimum of the energy function (equ. 13.5) with one to three constraints given by the plane equations (equ. 13.6).

$$f(\mathbf{p}) = \frac{1}{2}(\mathbf{p} - \mathbf{q})^2 \tag{13.5}$$

$$E_i : \ \mathbf{n_i p} - d_i = 0 \tag{13.6}$$

By the method of LAGRANGE-multipliers the function to be minimized can be set up as

$$h(\mathbf{p},\lambda_1,\lambda_2,\lambda_3) = \frac{1}{2}(\mathbf{p}-\mathbf{q})^2 + \sum_{i=1}^{3}\lambda_i(\mathbf{n_i}\mathbf{p}-d_i) \tag{13.7}$$

The minimum is found by setting the partial derivatives $\frac{\delta h}{\delta p_i}$ and $\frac{\delta h}{\delta p_i}$ to zero and solving the resulting system of equations:

$$\begin{pmatrix} 1 & 0 & 0 & n_{10} & n_{20} & n_{30} \\ 0 & 1 & 0 & n_{11} & n_{21} & n_{31} \\ 0 & 0 & 1 & n_{12} & n_{22} & n_{32} \\ n_{10} & n_{11} & n_{12} & 0 & 0 & 0 \\ n_{20} & n_{21} & n_{22} & 0 & 0 & 0 \\ n_{30} & n_{31} & n_{32} & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} p_0 \\ p_1 \\ p_2 \\ \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} = \begin{pmatrix} q_0 \\ q_1 \\ q_2 \\ d_1 \\ d_2 \\ d_3 \end{pmatrix} \tag{13.8}$$

The matrix is symmetric and always contains the identity matrix within the upper left 3×3 submatrix and the zero matrix within the lower right corner. In the first two passes per time step only one respectively two constraints are defined. Consequently the system of equations is reduced to the upper left 4×4 respectively 5×5 matrices. In the third pass it is sufficient to solve the system of equations according to

$$\begin{pmatrix} n_{10} & n_{11} & n_{12} \\ n_{20} & n_{21} & n_{22} \\ n_{30} & n_{31} & n_{32} \end{pmatrix} \begin{pmatrix} p_0 \\ p_1 \\ p_2 \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix} \tag{13.9}$$

which represents the calculation of the intersection points of the three supporting planes, as the solution for the Lagrange-multipliers is of no importance.

Due to the very simple structure all systems of equations can be inverted analytically in all three cases so that just the coefficients have to be set in at runtime. It is recommended to derive manually optimized program code out of this method, minimizing the number of necessary mathematical operations. As an example, the solution of the 3×3 system of equations (equ. 13.9) can be implemented most efficiently with the CRAMER's rule[23].

An extension of the god-object algorithm from a one-point interaction to an interaction with a spherical probe was suggested by RUSPINI, KOLAROV and KHATIB [211] in 1997. It solves one problem of the god-object algorithm that due to numerical errors the haptic probe may drop at edges of polygons into the inner volume of the object. By replacing the dimensionless point by an object with volume this effect of numerical gaps can be prevented.

Additionally the algorithm, which in some haptic-libraries implementing it is called RUSPINI algorithm, contains a concept for *force-shading*. It is used for smoothing the sensibly edged surface of polygonal objects resulting from the discontinuities of the normal directions at vertices and edges.

In graphical rendering e.g. PHONG-*shading* is used to cover the edges between polygons. For lighting calculation for each surface point a normal vector is used which is interpolated from the normal vectors stored at each vertex of the polygon.

[23] For larger systems of equations this approach leads to inefficient code though.

This idea is adapted for force-shading, but has to be extended by an additional step. As the direction of the force is defined by the direction of the spring-damper system between the haptic probe and the HIP, a new position for the haptic probe has to be derived from the interpolated normal first. The authors of the algorithm propose an adaptation in two phases: after calculating a first position $\mathbf{p}$ according to the constraint-method of the god-object algorithm, a plane is put through $\mathbf{p}$ whose normal vector is collinear to the interpolated normal. Onto this force-shading plane the HIP $\mathbf{q}$ is projected. In the second phase the final position $\mathbf{p}'$ of the haptic probe is calculated by using the newly found point $\mathbf{q}'$ in a second pass of the constraint-method. By this means the direction of the spring-damper system approximates the the direction of the interpolated normals.

An interesting fact about this approach for smoothing the direction of forces is the modification of the displayed force, which is achieved by altering the position of the haptic probe, and not by directly adding forces to the HIP. Avoiding a conceptional break in doing so is highly valuable for the stability of the method. Additionally this method is suitable for the integration of friction, the display of locally different stiffness of the same object, and the generation of fine surface structures by *haptic textures*.

13.3.4 6 DoF-interaction: Voxmap-PointShell-Algorithm

Another fundamental algorithm was developed by McNeely, Puter-Baugh and Troy [169] from Boeing in 1997 to solve questions in the area of virtual prototyping. The VPS algorithm combines a penalty-based approach including hybrid collision-detection with the simulation of rigid-body dynamics and the concept of a *virtual coupling* as suggested in 1993 by Colgate et al. [40] (see below).

In contrast to the methods discussed above the algorithm allows the output of a force vector and additionally of torques in three degrees of freedom. This is a basic requirement for the interactive planning of assembly pathes with virtual prototypes. It makes use of the tool-handle-metaphor and equates the virtual part with the handle of the haptic device.

One challenge in the development of the algorithm was to enable haptic rendering for highly complex scenes with huge polygon counts as they are generated from airplane CAD data. A clever representation of the scene was the right solution: the collision detection is not performed on the polygonal data but in form of hybrid collision detection between a so called *Voxmap* representing the static part of the scene and a so called *PointShell* which represents the “dynamic object”, i.e. the part to be assembled.

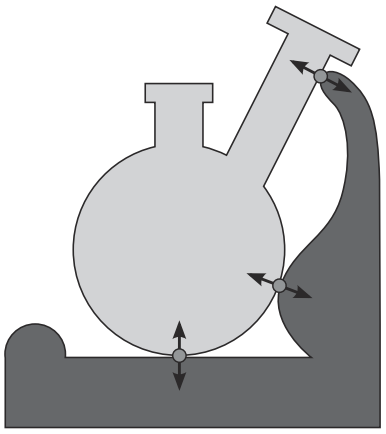


Fig. 13.7 Collision between dynamic object (light grey) and static part of scene (dark grey) with anti-parallel surface normals at the contact points.

Hybrid Collision Detection

The Voxmap is a global data structure representing the whole static scene. It consists of a three-dimensional grid of cubes and is generated in a preprocess in several steps. In each grid element or *Voxel* (abbreviation for volume-element in analogy to pixel) the information is stored, whether it is part of an object, whether it contains surfaces or whether it represents empty space of the scene (fig. 13.8). In a following step the voxels are classified in more detail, by e.g. storing the distance to the next surface point in each empty space voxel[24]. Afterwards the Voxmap is available as three-dimensional array of scalar data in a coherent memory area.

The rasterization process of the first generation step for the volume generation can be performed highly efficient on the graphics card by simply rendering the scene layer for layer. The volume model is the stack of the generated layers. The resolution of the layers should be chosen according to the detail level necessary for the application. As an example: a resolution of 1 mm is usually sufficient for assembly in automotive applications.

The dynamic part, the one to be moved by the user through the scene, is represented by the PointShell (fig. 13.9), which can be calculated in advance too. It consists of a multitude of points located at the surface of the dynamic object comparable to the data a laser-scanner would generate. Also the digital generation process of the point cloud is similar to a scan. The points are sampled by ray-tests in a regular raster, which can be realized with the help of the graphics card too. Once again the density of the sampling should be chosen to resemble the main details of the object and has to be tuned to the resolution of the Voxmap. For each surface point

[24] The calculation of this distance-field could be done via an *Euclidian-distance-transformation* (EDT)

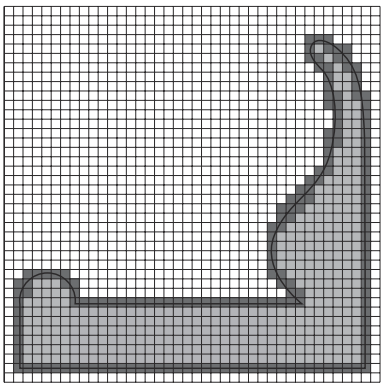


Fig. 13.8 Voxmap of the static part of the scene: surface-voxels are dark grey, inner voxels light grey and free-space voxels are white.

the corresponding surface-normal is calculated and stored negated. The reason for this is given a little further below.

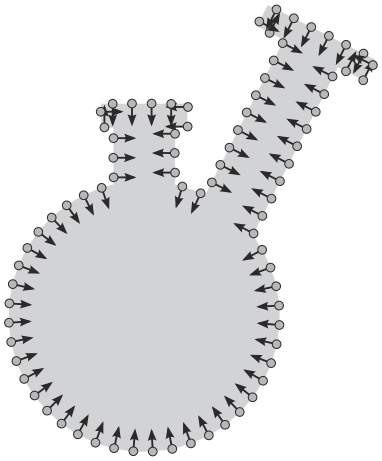


Fig. 13.9 PointShell of the dynamic object with negated surface normals

Within each cycle of the haptic rendering a simple collision detection algorithm is applied on the Voxmap-PointShell-representation of the scene: according to the actual transformation of the dynamic object, each point of the PointShell is transformed into the coordinate system of the Voxmap. Afterwards it is checked, in which voxel the point comes to rest: if the hidden voxel represents empty space, no local collision happens. Otherwise the point must be taken into account in the subsequent

classification and force simulation. The advantage of this hybrid representation in the collision detection is given by the fact, that for each PointShell-point only one access to the Voxmap has to be done. In addition, this access could be calculated highly efficient directly from the transformed coordinates of the PointShell-point: the three-dimensional index into the Voxmap can be achieved by rounding the point-coordinates to integer values.

As a result the runtime behaviour of the collision detection is completely independent of the complexity of the static model part of the scene. This advantageous runtime complexity however is paid with relatively high memory footprint[25], as the memory requirements scale cubically with the resolution of the Voxmap.

The runtime complexity of the total collision-detection step with one test for each PointShell-point scales linearly with the number of points in the point cloud. The latter in itself depends quadratically on the size of the dynamic object in relation to the sampling density. Consequently, dynamic objects with a large surface may be problematical[26].

Local Penalty Method

After collecting all colliding points of the PointShell local penalty-forces $\mathbf{F}_i$ are calculated and summed up. For the identification of the individual forces the surface normals $\mathbf{n}_i$ stored in the PointShell will be used: during contact of two real objects their surface normals at the contact point are anti-parallel (fig. 13.7). On a microscopic size this holds also for edges and vertices. As a consequence the contact normal for each PointShell-point can be used as surface normal of the dynamic object too.

The next challenge is to identify the penetration depth of the PointShell-point $\mathbf{p}_i$. With a clever trick this problem is transferred to the model of the virtual wall. A plane is put into the center $\mathbf{m}_j$ of the surface-voxel containing $\mathbf{p}_i$ with the normal direction $\mathbf{n}_i$. Then, a distance vector $\mathbf{p}_i$ to this plane is calculated. Multiplied with a spring constant the local penalty force $\mathbf{F}_i$ (fig. 13.10 and equ. 13.3) is calculated:

$$\mathbf{F}_i = \begin{cases} K(\mathbf{m}_j\mathbf{n}_i - \mathbf{p}_i\mathbf{n}_i)\mathbf{n}_i & \text{for } \mathbf{p}_i\mathbf{n}_i < \mathbf{m}_j\mathbf{n}_i \\ 0 & \text{else} \end{cases} \tag{13.10}$$

If the local forces $\mathbf{F}_i$ of all n detected contacts have been calculated, the total collision force $\mathbf{F}_K$ and the total collision torque $\mathbf{M}_K$ is summed up:

[25] A quite common correlation: the more memory can be used, the better the runtime performance will be.

[26] In theses cases hierarchic bounding volume methods (section 13.4) provide some room for optimization, which however are not discussed further in the context of this book.

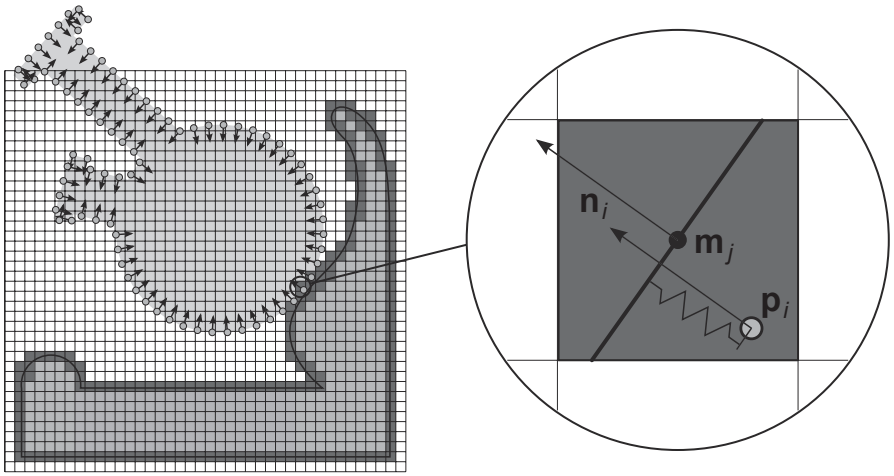


Fig. 13.10 Determination of the local penalty force

$$\begin{aligned} \mathbf{F}_K &= \frac{1}{n}\sum_i \mathbf{F}_i \\ \mathbf{M}_K &= \frac{1}{n}\sum_i \mathbf{F}_i \times \mathbf{l}_i \end{aligned} \tag{13.11}$$

with $\mathbf{l}_i = \mathbf{p}_i - \mathbf{S}$ and $\mathbf{S}$ being the actual position of the reference point (preferably the center of gravity) of the dynamic object.

Scaling by the factor $1/n$ avoids the emergence of too strong forces and smoothes the effects of contact-points being added and falling away between the cycles.

Now an obvious idea would be to actuate the 6 DoF haptic device directly with F_k and M_k. This would be identical to a hard coupling between HIP and the dynamic object resembling the haptic probe. Practically this approach frequently results into instable system behavior and a bad quality of the haptic impression. The dynamics of the haptic device is usually not sufficient to display the simulated forces.

Virtual Coupling

Similar to the god-object algorithm, which derives the force to be displayed by the device from the force of the spring-damper system between haptic probe and HIP, a so called *virtual coupling* couples the dynamic object with the handle of the haptic device. The position of the handle in the virtual scene is given by the HIP. In the VPS algorithm the coupling is designed as a six DoF spring-damper system: translatory springs with corresponding dampers represent three translatory DoF and rotatory springs with corresponding dampers the other three rotatory DoF. The simulation

of the torques generated by the virtual coupling is implemented analogously to the force-simulation of the translatory spring-damper system:

$$\begin{aligned} \mathbf{F}_V &= K_t \mathbf{q} - D_t \mathbf{v} \\ \mathbf{M}_V &= K_r \theta - D_r \omega \end{aligned} \tag{13.12}$$

with K_t and K_r beeing translatory and rotatory spring constants and D_t and D_r representing the corresponding damping constants. The values $\mathbf{q}$ and $\mathbf{v}$ are the vector distance between dynamic object and virtual handle (representing the displacement of the virtual spring in idle position 0) and their *relative* velocity. The vector θ represents the relative rotation between the dynamic object and the virtual handle, ω is the relative angular velocity.

Rigid-Body Dynamics

Up to now it is open, how the position and motion of the dynamic object is identified within the virtual scenario. In the steps described till here each haptic rendering cycle calculates forces and torques of the virtual collisions and inside the virtual coupling. Now still their impact on the dynamic object has to be simulated.

One approach is given by the integration of motion equations according to NEWTON and EULER[27]:

$$\begin{aligned} \mathbf{F} &= \mathbf{F}_K + \mathbf{F}_V = m\mathbf{a} \\ \mathbf{M} &= \mathbf{M}_K + \mathbf{M}_V = \mathbf{I}\alpha + \omega \times \mathbf{I}\omega \end{aligned} \tag{13.13}$$

with the translatory acceleration $\mathbf{a} = \dot{\mathbf{v}}$ and the angular acceleration $\alpha = \dot{\omega}$. As additional necessary values the mass m of the dynamic object and its inertial tensor $\mathbf{I}$ describing the mass distribution of the object have to be provided.

Both values can be derived from the surface model of the object with the methods already implemented: a volume model of he dynamic object is generated with the same voxelization-method known from the generation of Voxmaps. To each center $\mathbf{w}_k$ of an occupied voxel a point mass m_k is assigned. The total mass m is given by the sum of all m_k and the components I_{ik} of the inertial tensor

$$I = \begin{pmatrix} I_{00} & I_{01} & I_{02} \\ I_{10} & I_{11} & I_{12} \\ I_{20} & I_{21} & I_{22} \end{pmatrix}$$

are calculated with

[27] For a very helpful introduction into the implementation of rigid-body dynamics simulation based on NEWTON-EULER equations [13] is recommended.

$$I_{ij} = \sum_{k=1}^{n} m_k(\mathbf{r}_k^2 \delta_{ij} - \mathbf{r}_{ki}\mathbf{r}_{kj}) \tag{13.14}$$

with the KRONECKER-symbol δ_{ij} and $\mathbf{r} = \mathbf{v} - \mathbf{S}$. $\mathbf{S}$ denotes the point of action of the virtual coupling at the object. Optimally the object is modelled so that $\mathbf{S}$ coincides with the center of mass of the object. The latter is important to prevent the emergence of torques during translatory movements in collision-free space, which otherwise would disturb the haptic impression.

Including these components the VPS algorithm is complete: in the virtual scene the collision detection and identification of collision-forces is done. Together with the forces and torques from the virtual coupling they act upon the dynamic object and result in its movement. Finally the torques and forces of the virtual coupling are displayed by the haptic device. The coupling therefore builds the bridge between real handle and virtual part: it is like the user would move the dynamic object through the world by pulling a rubber band attached to it. According to the application the haptic impression can be controlled by configuring the mass of the dynamic object adapted to the spring stiffness and particularly to the damping of the virtual coupling.

This overview about the VPS algorithm shows how object representations, collision detection, physical modelling and finally the simulation of the forces to be displayed are interconnected and have to be harmonized with each other.

At the same time the limits of a certain modelling concept for a usage beyond the intended application become obvious. A volumetric description of the static part of the scene cannot be used without major efforts to represent deformable models. Movements of single objects of the static scene or interaction between objects except for the dynamic object cannot be realized directly with the hybrid collision detection. Nevertheless it should be noted, that the algorithm proved very valuable for the interactive verification of assembly steps of rigid parts in rigid installation spaces. For this purpose it is one of the best methods at hand.

13.4 Collision Detection

The core element of a haptic-rendering algorithm is efficient collision detection. At a frequency of close to 1000 Hz it must be checked whether there is contact between the haptic probe and the virtual object. Collision detection is also relevant for applications in the area of virtual-reality which not necessarily include haptics. Fundamentally any kind of physics simulation, where virtual objects interact with each other, should be mentioned. In scientific as well as in industrial applications and in computer games the collision detection is of central importance. Consequently it was focused on this topic since the very beginning of virtual reality.

Beside the obvious usages of collision detection algorithms — to find out whether virtual objects collide — there are other less obvious ones: within so called

raytracers the paths of single light rays are simulated physically to create a photo-realistic image. For one image millions of collision tests have to be made between rays and the virtual scene (using ray intersection tests).

Based on the following considerations the most fundamental questions and concepts of collision detection methods as well as principles for their optimization shall be discussed. This will provide a good basic understanding of the challenges related to it. For a more comprehensive discussion specialized literature for collision detection is recommended [295, 50].

For all problems where large amounts of data have to be processed efficiently, algorithms with low *complexity* have to be designed. Complexity is a measure how processing-time, memory requirements, and the need for other resources for the algorithm depend on the problem size, which is usually measured by the number of elements to be processed.

If for example a collision between two objects shall be tested, this can be done "brute force" by testing all triangles of the one object against all triangles of the other object. The effort increases quadratically with the size of the problem: if the number of triangles of both objects doubles, the runtime of the algorithm quadruples. Thus the algorithm has quadratic order, the (runtime-)complexity is $O(n^2)$.

The size of the models which can be displayed graphically on state-of-the-art VR-systems and even on off-the-shelf PCs in real time (section 13.2.4) is several hundred-thousand or even more than a million triangles. Even in such complex scenarios efficient collision detection is desired, quadratic complexity however is completely impracticable.

The total complexity of a collision detection method can be improved with a plausible idea: in several steps the size of the problem is reduced by performing simple and more "economic" tests, excluding as many elements as possible from the more complex tests. The necessary steps are usually arranged in a *collision-pipeline*, which is processed step by step.

For a fast exclusion of many elements their spatial coherence is analyzed and exploited. For example neighbouring triangles can be collected into a bounding object. If a test of the bounding object shows no intersection, all contained elements can be excluded from all further analysis.

Different geometry types are suitable as bounding objects: spheres, cubes, convex polyhedrons with defined orientation of surfaces (so called *discrete oriented polytopes*, DOP), to name the most frequently used ones. Usually only one bounding geometry type is used within the same collision detection step. A "good" bounding geometry can be tested efficiently for intersections and approximates the included elements closely, because this reduces the number of false positive tests. The quality of the approximation achievable with a bounding geometry depends highly on the shape of the bounded object. Figure 13.11 shows this effect for different bounding geometry types: spheres for example approximate elongated objects only badly. Consequently flexible bounding geometry types are advantageous for the approximation quality. However they require more effort for intersection tests. A good compromise are DOPs which reach a good approximation quality in combination with relatively simple intersection tests. But there is no universal optimum choice

of a bounding geometry type for every application, as the runtime performance at collision detection is dependent on many factors.

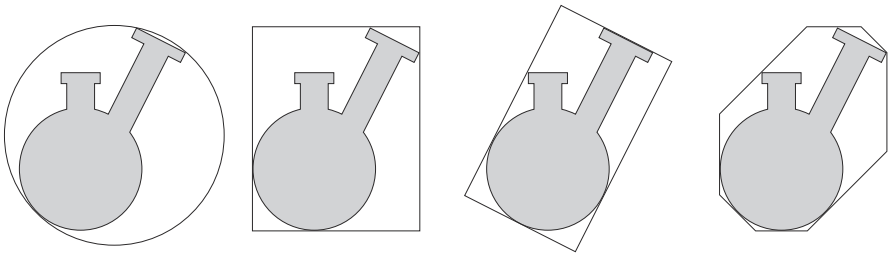


Fig. 13.11 Types of hull geometries in 2D (from left to right): minimal sphere, axis aligned bounding box (AABB), object oriented bounding box (OOBB), discrete oriented polytope (DOP)

The test for the intersection of two spheres is simple: may m_0 and m_1 be the centers of two spheres and r_0 and r_1 their radii. They do not intersect, if $|m_0 - m_1| > r_0 + r_1$. For an efficient implementation the monotony of the square in the first quadrant should be used to avoid the extraction of the root for the calculation of the distance between m_0 and m_1. The test condition then is $(m_0 - m_1)^2 > (r_0 + r_1)^2$. This saves one root for the costs of an additional multiplication for squaring the sum of the radii, which still requires considerable less processor cycles. The optimization of these tests is relevant, as they are frequently used within the collision detection process as depicted below. The calculation of the bounding sphere (which exists and is unique) is not trivial, but can be done using the algorithm of HOPP [96].

In average fewer effort is required by the test for intersection of two axis aligned boxes. One axis aligned box Q can be characterized by two vertices q_{min} and q_{max}. The intersection of both boxes is empty — no collision happens — if at least one coordinate of the minimum of one box is larger than the maximum of the other box (fig. 13.12).

This makes it possible to formulate a test, which has to be processed in the average of all collision-free cases only to the half of its code, as it already excludes a collision before:

```
if      qMin[0] > pMax[0]:
  return "no collision"
else if qMax[0] < pMin[0]:
  return "no collision"
else if qMin[1] > pMax[1]:
  return "no collision"
else if qMax[1] < pMin[1]:
  return "no collision"
else if qMin[2] > pMax[2]:
  return "no collision"
```

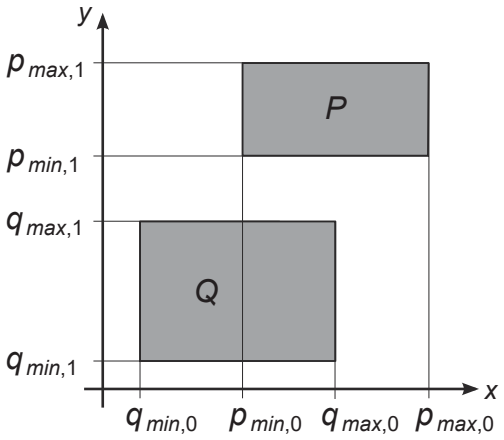


Fig. 13.12 Intersection test of boxes: no intersection of the intervals on thed *y*-axis guarantess absence of collisions.

```
else if qMax[2] < pMin[2]:
  return "no collision"
else:
  return "boxes intersect"
```

The full potential of tests via bounding volumes is used in hierarchic collision detection algorithms working with trees of nested bounding volumes[28].

The root of such a tree resembles the total object, which is nothing but “all triangles”. During generation of the bounding-volume hierarchy (BVH) all triangles of a node are distributed onto its child nodes recursively until the leaves of the tree contain exactly one triangle (fig. 13.13).

An advantageous split is essentially for the efficiency of the following collision detection. Quality criterions are

- The tree should be *balanced* according to the number of its child-nodes and therefore be optimized in its depth. This is identical to the requirement, that the geometric primitives of a node should be distributed equally among its child nodes.
- The distribution of the geometrical primitives should be chosen to minimize the volumes of the bounding geometries and reduce overlapping on the same level of the tree.

[28] A tree is a cycle-free interconnected *graph*, consisting of *nodes* and connecting *edges*. *Directed trees* have one *root*. *Leaves* are nodes which do not have any “child-nodes” and no outgoing *branches*. Usually hierarchic collision detection algorithms use binary trees, i.e. trees where all nodes except for the leaves have two child-nodes. *Octrees*, trees with eight child-nodes, are another frequent variation.

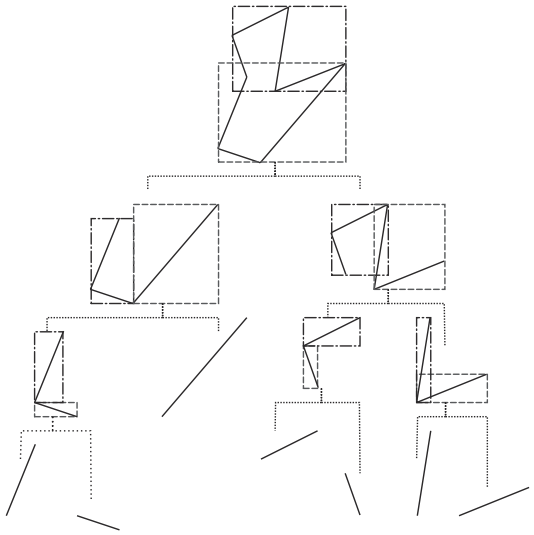


Fig. 13.13 Schematic representation of the split of a geometry inside a bounding-volume hierarchy

How good both criterions can be fulfilled at the same time depends on the structure of the individual models. For models made of triangles with large variations in their size, no optimum distribution can be found covering both criterions equally good.

After the generation of the bounding volume hierarchy in a preprocess the data structures are available, which enable the application to quickly filter out large parts of the objects during the collision tests. If the bounding volumes do not intersect in a node of the tree, all child nodes and the included triangles can be neglected for the following tests, whereas nodes, which could not be excluded, are recursively processed in more detail during the subsequent tests. Very fast only branches close to the actual contact point remain.

If the question is about the existence of collision only, the collision detection can be interrupted as soon as a collision was found for one branch. For this question therefore a near-contact leads to the worst average performance: the hierarchy has to be traversed deeply up to the demanding triangle tests to just return “no collision” as result finally.

However for haptic applications exact information about the contact is necessary to be able to classify it. This requires all contact points to be identified, i.e. all triangles which intersect another object. In this *exact* collision detection the near-collision doesn’t differ from collision concerning its runtime performance.

In many applications it is not sufficient to regard the collision between a geometrically simple probe (e.g. a point or sphere) with objects of the virtual scene but the contact between complex geometrical models. In this situation both tested objects are represented by bounding volume hierarchies, which are processed in parallel recursive traversal.

```
function CollisionTest( node0, node1 ):
  if bounding volumes of node0 and node1 are disjoint:
    return
  else if node0 and node1 are leafs:
    if TriangleTest( node0, node1 ) is positive:
      add triangle pair to collision list
  else if node0 "larger than" node1:
    for each child node cn of node0:
      CollisionTest( cn, node1 )
  else:
    for each child node cn of node1:
      CollisionTest( node0, cn )
```

The function `CollisionTest` of the pseudo-code-listing above is called with the root nodes of both object's bounding-volume hierarchies. During the recursive processing the list is filled with all pairs of triangles of the one or the other object which collide (it is an algorithm of the class of exact collision detections). If the list is empty at its end, in the current time-step there are no collisions. As metric (for the test "larger than") for the choice of the path to descend the number of children or the size of the bounding volume may be chosen.

Beside spatial coherence also temporal coherence is exploited often: in a system simulated with small time-steps usually only small changes happen from step to step. Information from the last collision detection cycle can be recycled as a starting point for the search of a collision in the current time-step. Collision detection algorithms according to this principle are called *incremental*.

A good example for this is an algorithm which often very quickly proves that no collision exists and thus used in early phases of collision pipelines. The convex hulls of two objects given it tries to identify a plane, which lies in between them. If so, the objects are in different of the both disjunctive half spaces generated by the plane. The existence of such a *separating plane* guarantees that no collision happens.

The first test within each cycle of the collision detection checks whether a separating plane existed within the prior cycle. This plane is tested for being valid in the current cycle too. If this is the case, the further processing of the collision pipeline for the object pair can be stopped. If this is not the case, a new separating plane is searched nearby the old plane until one is found or a maximum number of trials was made. In the latter case the processing of the collision pipeline is continued, as collision can't be excluded (fig. 13.14).

Another incremental procedure is given by the *Closest-Feature-Tracking* algorithm of LIN and CANNY [163]. It searches for the pair of surface features (points, edges, faces) of two polygonal objects which constitute the minimum distance between both (fig. 13.15). The method starts with an arbitrary feature pair and analyzes specificly[29] pairs of neighbouring features whether they are closer to each other. The

[29] For the choice of the next feature pair the VORONOI regions of the objects are analyzed.

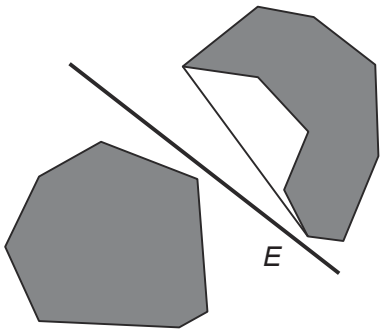


Fig. 13.14 Separating plane *E*

search gives a fast result if a good starting pair was chosen. Such a starting pair is often available in form of the feature pair of the last collision test.

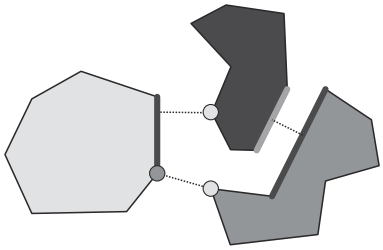


Fig. 13.15 Pairwise closest features of a set of objects

For the haptic rendering with a point-like probe the question for the collision detection has to be formulated different from the collision with complex models: it has to be identified whether the point is within the object, or whether the movement of the point passed a surface of an object within the last time-step.

A point is within a closed polyhedron[30] if a ray casted to infinity starting from this point in an arbitrary direction hits an odd number of surfaces. With convex polyhedrons the ray hits exactly once. With concave objects numerous hits may occur (fig. 13.16). When implementing this test some exceptional cases have to be considered: the ray may hit an edge or point within the numerical precision. In this case the corresponding surfaces have to be counted very carefully. Grazing touches of the corresponding surfaces (only relevant for the concave case) should not be counted. A penetration should be counted only once even when there are two or more surfaces touching the ray.

[30] An object modelled of triangles representing a solid body is such a closed polyhedron. For non-closed polyhedrons with a “hole” in their surface an inner region can not be defined so easily.

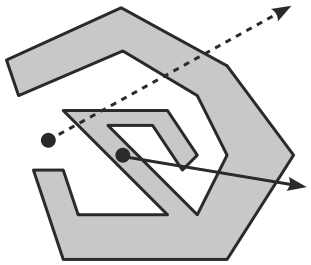


Fig. 13.16 Ray intersection tests with odd (drawn through ray) and even (dashed ray) number of intersection points for an inner and an outer point respectively.

The question about a point shape probe having crossed a surface within the last time-step is answered by a ray intersection test too, but the test is performed with the line between the last and the current position of the probe instead with the ray.

In both cases the ray intersection test can be optimized using the efficient collision detection methods described above: the ray resp. the line may be tested recursively against the bounding volume hierarchy of the object in question.

The analysis of the two collision queries for the one-point-interaction leads to another basic distinction. The collision tests between complex objects as described so far only consider static states of the objects within one time-step. As a consequence of the time discretization collisions can be missed if a relative small or fast object is in collision only in between two simulation steps. These cases are addressed by the methods of *dynamic collision detection* (in contrast to *static collision detection*), which additionally are suitable to calculate the moment of collision back in time and interpolate the corresponding position of the involved objects.

The basic concept is to approximate the volume occupied by the moving object since the last simulation step and to test this volume for collisions with other objects. It is important to avoid the calculation of the envelopes of *both* objects. If the relative movement of the objects is expressed within the coordinate system of one of the objects, it is sufficient to calculate the envelope of the other "dynamic" object.

The methods for hierarchic collision detection can be applied to dynamic collision detection too. We start from the assumption that for both involved objects boundary hierarchies exists. During the recursive collision detection algorithm an envelope for the bounding volume of the current node in the bounding volume hierarchy of the dynamic object is approximated and tested for intersection. Once again the principle is to exclude large parts of the geometry as effortless and early as possible from the more performance-critical collision tests.

The approximation of the envelope of the bounding geometry as well as of the individual triangles is simple for translations. In these cases the movement volume itself can be characterized by a polygonal object. The bounding surfaces of rotating objects become so called *ruled surfaces* though which are curved and need a more complex representation. Therefore once again a rough outer approximation of the envelope by an bounding object of the chosen bounding geometry type (sphere, box,

DOP) is used. Only in the leaves of the hierarchy exact tests are performed for the triangles, which are necessary when the contact moments are calculated back in time.

13.5 Conclusion

For the integration of haptic rendering in existing virtual-reality systems a modification of their architecture is usually necessary. In case of a new development of a multimodal system which should include the haptic sense an appropriate architecture should be chosen: it should provide functionalities for multithreading and/or distribution including mechanisms for synchronization. This is needed to decouple the higher-frequent haptic rendering of the application behavior and graphic rendering cycle.

Additionally the VR-system has to include data-structures and algorithms for exact collision detection between haptic probes and virtual objects as well as physical models for the calculation of collision- or frictional forces. If the forces calculated inside the haptic-simulation are intended to affect the virtual scene, instances for physical simulation, rigid-body dynamics, deformable objects etc. have to be implemented in the scenario and linked to the objects of the scene. The resulting data flow requires an open eye on the synchronization between the different modules.

Due to the hybrid functionality of haptic devices — they are input- and output-devices at the same time — an adequate concept of device abstraction has to be realized too. At least the communication between high-frequency haptic rendering thread and haptic device has to be designed with little latency. With “dangerous” devices, which are extremely strong or fast, it is strongly recommended to give hard realtime-guarantees to the haptic rendering loop. Among other things this requires to build on top of realtime operating systems. Softer realtime-requirements should be fulfilled for smaller haptic devices in any case, as the fundamental system stability and an acceptable quality of the haptic feedback can be achieved easier.

"*This page left intentionally blank.*"

Chapter 14
Final Remarks on the Design of Haptic Systems

Like any other design process the design of haptic systems is largely influenced by the optimization of a technical system based on the balancing of a plurality of decisions on separate components, which, as a rule, influence each other. In the beginning, the requirements of the customer respectively of the project have to be defined. The methods presented in chapter 6 are intended to systematically identify the most important aspects of these requirements. However, the engineer should be conscious of the fact that for the design of a sense-related interface less precise and definite terms are available than he may be used to. Additionally, the knowledge on the part of the customer may result in considerable confusion, as especially haptic terms like, e.g., resolution or dynamics may be used in the wrong context or understood in a wrong way. A better definition of the requirements without major misunderstandings is achieved by e.g. giving the customer aids, "shows-and-tells" of haptics. It is necessary for the customer and the engineer to come to a common understanding based on references known to both.

For this reason, an understanding of the specialties of haptic perception on the part of the engineer is necessary. It should not be limited to the technical characteristics described in chapters 2 and 3, but also include some knowledge about the "soft", i.e. psychological and social aspects of haptics (chapter 1). In chapter 4 an attempt is made to bridge the gap between these aspects and the technical description of haptic perception. However, the application of the methods presented requires increased attention, as the data basis of the models is not yet sufficient. Nevertheless, they show a path suitable for further optimization of haptic devices.

Based on the requirements discussed above, the technical design process may begin. One of the very first decisions is the choice of the haptic system's structure (chapter 5). Although this decision is at the very beginning of the design process, a rough sketch of the favored structure of the device to be developed has necessarily

T.A. Kern (ed.), *Engineering Haptic Devices*, DOI 10.1007/978-3-540-88248-0_14, 431
© Springer-Verlag Berlin Heidelberg 2009

to be made. This demands a considerable previous knowledge of all the branches of haptic device design, which later will be needed again during the actual design phase. Besides the already mentioned decision on the general structure (chapter 7) , the basis of the design of kinaesthetic and tactile systems is its kinematic structure (chapter 8). After the considerations made for kinematics, concerning the transmission- and gearing-proportions, the working volume, and the resolution to be achieved, suitable actuators are chosen or even designed. In chapter 9 the basis for this is provided by comparing the different actuation principles. Examples of their realizations, even of unusual solutions for haptic applications, provide a useful collection for any engineer to combine kinematic requirements of maximum forces and translations with impedances and resolutions.

As closed-loop admittance controlled systems with kinaesthetic and tactile applications are gaining in importance, force sensors have to be considered as another component of haptic devices. In chapter 10 this technology is introduced, providing the tools, as well as conveying the chances but also the challenges connected with their application.
A frequent application of haptic devices is to be found in the human-machine interface of simulators, be it for games ranging from action to adventure games, or for more serious applications for training surgeons or in the military, respectively in industrial design. In addition to the output of haptic information an input of user-movements is required. The measurement principles typically used are discussed in chapter 11. The design steps presented so far will enable the haptic device to provide a tactile or kinaesthetic output to the user, often measuring a reaction, too. Especially with today's computer technology, the data will be almost always interfaced with a standard PC. The requirements derived from this interface are subject to a presentation of standard interface technology given in chapter 12, whereby the interfaces' performances are compared with each other.

Due to the rather frequent application of haptic devices in simulators, an interface with a simulation engine is required. An insight into the requirements and challenges of suitable haptic algorithms is helpful for the hardware engineer to improve the communication with software-engineers and the interfacing with their VR-environments. An appropriate introduction is given in chapter 13.

The cross-section given in this book is meant to improve and further speed up the design of haptic devices, and to avoid the most critical errors typically made during the design process. The research on the area of haptic devices is making impressive progress. Every few months adapted control-engineering concepts appear; the usage of haptic perception for the design is subject to current research. Actuators are being continuously improved; even new principles with haptically interesting properties regularly appear on the market. Closed-loop controlled systems become more and more interesting, due to the slowly increasing availability of highly dynamic high-resolution force sensors, This dynamics of a still young discipline commits the developing engineers to monitor current research attentively. For this purpose,

finally, a list enumerating teams active in the haptic area has been compiled in the appendix-chapter 15.

"*This page left intentionally blank.*"

Part III
Appendix

"*This page left intentionally blank.*"

Chapter 15
URLs

The following tables are a collection of all names and groups familiar to the authors. Naturally the lists are not compelling, but they provide a starting point for the research about relevant sources. Additions to this list for future editions of this book may be sent to the editor via email.

Table 15.1: URLs of labs and individuals working in the area of haptics, taken and extended from HapticsSymposium.org

Institute	Head	URL	Type
ALAB, Tokyo Univ.	SHINODA	www.alab.t.u-tokyo.ac.jp/	Lab
A Labaratory for Teleoperation and Autonomous Intelligent Robots (ALTAIR)	FIORINI	http://metropolis.sci.univr.it/r_d/r_d.php	Lab
Artificial Intelligence Lab, Robotics	KHATIB	http://robotics.stanford.edu/	Lab
Bionengineering, Department of	BURDET	www.bg.ic.ac.uk/staff/burdet/Home.html	Person
Bioinstrumentation Lab., Haptic group	JONES	http://bioinstrumentation.mit.edu/jones/	Person
Biorobotics Labaratory	HANNAFORD	http://brl.ee.washington.edu	Lab
BioRobotics Labaratory	RYU	http://robot.kut.ac.kr/	Lab
Biomimetics and Dexterous Manipulation Labaratory	CUTKOSKY	http://bdml.stanford.edu/DML	Lab
Computer Graphics Labaratory ETH Zürich	GROSS	http://graphics.ethz.ch	Lab
Delft Haptics Lab	HELM and others	www.tudelft.nl	Lab

see next page ...

T.A. Kern (ed.), *Engineering Haptic Devices*, DOI 10.1007/978-3-540-88248-0_15,
© Springer-Verlag Berlin Heidelberg 2009

... continuation

Institute	Head	URL	Type
EduHaptics	PROVANCHER	www.eduhaptics.org	Edu.
Fujimoto Lab	FUJIMOTO	http://drei.mech.nitech.ac.jp/ fujimoto/	Lab
Group of Robots and Intelligence Machines	FERRE and others	http://138.100.76.36/en/default.asp	Group
Haptics and Embedded Mechatronics Lab	PROVANCHER	http://heml.eng.utah.edu/pmwiki.php/Main/HomePage	Lab
Haptics Grasp Lab	KUCHENBECKER	http://haptics.grasp.upenn.edu/	Lab
Haptic Exploration Lab	OKAMURA	www.haptics.me.jhu.edu/	Lab
Haptic Interface Research Lab	TAN	www.ecn.purdue.edu/HIRL/index	Lab
Haptics Research Group	PANCHANATAN	http://haptics.asu.edu/index.php	Lab
Haptics and Virtual Reality Lab	CHOI	http://hvr.postech.ac.kr/wiki/wiki.php	Lab
Haptics Labaratory	HAYWARD	www.cim.mcgill.ca/%7Ehaptic/grouphome.html	Lab
Haptiklabor	HEGEL	www.haptiklabor.de/	Lab
Haptix Labaratory	GILLESPIE	http://www-personal.umich.edu/%7Ebrentg/Web	Lab
Harvard Biorobotics Labaratory	HOWE	http://biorobotics.harvard.edu/	Lab
Human Factors Engineer	ANDREW	andyand@talktalk.net	Person
Human machine interface Lab	BURDEA	www.caip.rutgers.edu/vrlab/	Lab
Human sciences group, Insitute of sound and vibration reserach	GRIFFIN	www.isvr.soton.ac.uk/HSG	Inst.
Institute of Automatic Controle Engineering	BUSS, HIRCHE, STURSBERG, SCHMIDT	www.lsr.ei.tum.de/	Inst.
Institute of Robotics and Mechatronics	HIRZINGER	www.dlr.de/rm/en/	Inst.
Institute for Elektromechanical Design	SCHLAAK, WERTH-SCHUETZKY	www.institute-emk.de/	Inst.
Interactive Systems Research Group (ISRG)	HARWIN	http://www.reading.ac.uk/isrg/isrg-haptics.asp	Group
Interdisciplinary Institute for Biomedical Research (IINR)	HOWE	www.oucom.ohiou.edu/IINR/	Lab
Ishibashi and Sugawara Lab	ISHIBASHI, SUGAWARA	http://nma.elcom.nitech.ac.jp/index_e.html	Lab
Laboratoire de Systèmes Robotiques (LSRO)	BLEULER	http://lsro.epfl.ch	Lab
Laboratoire Systemes Complexes (Fre-CNRS)	KHEDDAR	http://lsc.univ-evry.fr/	Lab
Mechatronics and Haptics Interfaces (MAHI)	O'MALLEY	http://mems.rice.edu/ %7Emahil	Lab
Mechatronics Lab	YOKOKOHJI	www.mechatronics.me.kyoto-u.ac.jp	Lab

see next page ...

... continuation

Institute	Head	URL	Type
Microdynamic System Labaratory	HOLLIS	www.msl.ri.cmu.edu/	Lab
MIT Touch Lab	SRINIVASAN	http://touchlab.mit.edu/	Lab
Multimedia Communication and Research (MCR) Lab	SADDIK and others	http://www.mcrlab.uottawa.ca/	Lab
Multimodal Interaction Group	BREWSTER	www.dcs.gla.ac.uk/ stephen/	Group
Multisensory Perception and Action	ERNST	www.kyb.tuebingen.mpg.de/ bu/people/-marc	Lab
PERCRO	Bergamasco	www.percro.org/	Lab
Physics of man, department	KAPPERS, KOENDERING	www.phys.uu.nl/ wwwpm/HumPerc	Inst.
Precision and Intelligence Laboratory	SATO	http://kenwww.pi.titech.ac.jp/	Lab
Psychology at Hamilton	GESCHEIDER, BURR, YEE, WELDON, VAUGHAN	http://academics.hamilton.edu/psychology/home/	Inst.
Psychology at Carnegie Mellon, spacial and haptic	KLATZKY	www.psy.cmu.edu/faculty/klatzky/lab/	Lab
Labaratories for Intelligent Machine Systems (LIMS)	COLGATE, PESHKIN, LYNCH	http://lims.mech.northwestern.edu	Labs
Robotics Group, "E.Piaggio"	BICCHI	www.piaggio.ccii.unipi.it/newrobotics/ roboticresearch.html	Group
Salisbury Research Group	SALISBURY	http://jks-folks.stanford.edu/home.html	Lab
Sensory Motor Neuroscience	WING	www.symon.bham.ac.uk/labs.htm	Lab
Sensory Perception and Research Group	MACLEAN	www.cs.ubc.ca/labs/spin	Lab
Skripte zur Psychologie	ZWISLER	www.zwisler.de/	Person
Tachi Lab	TACHI	http://tachilab.org	Lab
Telerobotics and Control Lab (TCL)	KWON	http://robot.kaist.ac.kr/	Lab
Telerobotics Lab	NIEMEYER	http://telerobotics.stanford.edu	Lab
The Senses of Touch	PATTERSON	www.ggy.bris.ac.uk/postgraduates/ggmp	Person
TNO, Haptics	ERP, VAN	www.tno.nl	Inst.
Touch Laboratory	LEDERMAN	http://psyc.queensu.ca/%7Echeryl/labpage.html	Lab
Haptic group Utah	HOLLERBACH, JOHNSON, MASCARO, PROVANCHER	www.cs.utah.edu/ jmh/Haptics.html	Group
Virgina Touch Laboratory	GERLING	www.sys.virginia.edu/ggerling/	Lab

see next page ...

... continuation

Institute	Head	URL	Type
Virtual Reality and Active Interfaces (VRAI) Group	BAUR	http://vrai-group.epfl.ch/	Group
Virtual Reality in Medicine	HARDERS	http://www.vision.ee.ethz.ch/research/projects _med.cgi?topic=6	Lab
Virtual Reality Lab	KESAVADAS	www.haptics.buffalo.edu	Lab
Visualisation and Image Analysis Labaratory (VIALAB)	STETTEN	www.vialab.org	Lab
VR-Lab. Univ. of Tsukuba	IWATA, YANO	http://intron.kz.tsukuba.ac.jp/index_e.html	Lab

Table 15.2 Established conferences, workgroups, journals with reasonable parts of haptic hardware

Title	URL	Type
Eurohaptics Conference	www.eurohaptics.com	Conf.
Haptics-e	www.haptics-e.org/	Journ.
HapticsSymposium	www.hapticssymposium.org	Conf.
Haptics technical comittee (IEEE)	www.worldhaptics.com	Soc.
Transactions on Haptics (IEEE)	www.computer.org/th	Journ.
Worldhaptics Conference	www.worldhaptics.com	Conf.

Table 15.3 Commercial manufacturers of haptic-related products

Manufacturer	URL	Type
Artificial Muscle	www.artificalmuscle.com	HW
Butterfly Haptics	www.butterflyhaptics.com	HW
Chai3D	www.chai3d.org	SW
ForceDimension	www.forcedimension.com	HW
Guitammer Company Inc.	www.ButtKicker.com	HW
haptiklibary	www.haptiklibrary.org	SW
Haption	www.haption.com	HW
Haptx	www.haptx.com	SW
Immersion	www.immersion.com	HW & SW
Moog	www.fcs-cs.com/robotics	HW
MPB Technologies Inc.	www.mpb-technologies.ca	HW
Novint	www.novint.com	HW
Quanser	www.quanser.com	HW
reachin	www.reachin.se	SW
SensAble	www.sensable.com	HW & SW
Sensegraphics	www.sensegraphics.com	SW & HW
Xitact	www.xitact.com	HW

"*This page left intentionally blank.*"

Chapter 16
Mechanical Impedances and Admittances for Translatory and Rotatory Systems

THORSTEN A. KERN

For the description of technical systems a model of the real physical properties is mandatory. A central toolset in the area of electromechanics is the use of linear, time-invariant, and concentrated elements. The following section gives a short introduction into this topic. Further details can be found especially in [158].

Concentrated elements represent real physical object properties. They simplify these properties in three fundamental aspects:

- The objects represented by the elements do not have any spatial size - they are concentrated on a single spot.
- The object properties are linear.
- The object properties are time-independent.

Accepting above limits three types of elements are sufficient to model all relevant mechanical effects.:

- Mass/inertia m resp. moment of inertia Θ
- Linear or rotary spring with elasticity n or spring-constant/stiffness k
- Viscous friction resp. rotatory viscou friction with damping d resp. frictional coefficient r

They are represented by the symbols according to figure 16.1.

Using these elements modelling can be made in time- and frequency domain. For the analysis about stability and frequency plots a description in LAPLACE-domain with complex numbers is advantageous. As there is a limited operation time expected in all cases the complex frequency is given as $s = j\,\omega + \varphi$ (mit $j \cdot j = -1$). For calculating the complex system behaviour two quotients are distinguished:

- mechanical impedance $\underline{Z}_{\text{transl}} = \frac{\underline{F}}{\underline{v}}$ resp. $\underline{Z}_{\text{rot}} = \frac{\underline{M}}{\underline{\Omega}}$
- mechanical admittance $\underline{Y}_{\text{tranl}} = \frac{\underline{v}}{\underline{F}}$ bzw. $\underline{Y}_{\text{rot}} = \frac{\underline{\Omega}}{\underline{M}}$

T.A. Kern (ed.), *Engineering Haptic Devices*, DOI 10.1007/978-3-540-88248-0_16,

© Springer-Verlag Berlin Heidelberg 2009

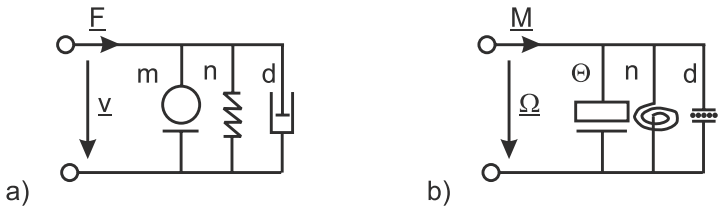


Fig. 16.1 Symbols for concentrated mechanical elements for translatory (a) and rotatory (b) systems according to the 2nd analogy.

For the elements this gives us:

translatory mass:	$s\,m = \underline{Z}$
translatory spring:	$\frac{1}{s\,n} = \frac{k}{s} = \underline{Z}$
translatory friction:	$r = d = \underline{Z}$
rotatory mass (moment of inertia):	$s\,\Theta = \underline{Z}$
rotatory spring:	$\frac{1}{s\,n} = \frac{k}{s} = \underline{Z}$
rotatory friction:	$r = s = \underline{Z}$

The elements can be combined to networks. Their usage happens analog to electrical networks. Depending whether velocity or force (torque or rotational velocity) is used as differential value, two analogies can be formulated. They differ according to the handling of the node- and mesh-equations in network-theory. This book makes use of the 2nd analogy, whereas $\underline{F}$ is the flux- and $\underline{v}$ the differential-value. A transition of the mechanical networks in their electrical equivalent circuit does not happen in the context of this book. However learning and using this method is recommend to any electrical engineer to improve the understanding of mechanical systems. The limitations of concentrated elements can be avoided by a transition of the network to finite elements and a solution of the resulting wave-equations. However in the context of this book this advanced method is not used.

Chapter 17
Details about Gyrators and Transformers

STEPHANIE SINDLINGER

For the design of electromechanical systems the interdependencies between electrical and mechanical networks are formulated by the use of transducers. These transducers are linear, lossless and time-invariant quadrupoles.

The electromechanical converters are distinguished in *active* and *passive* ones. They show the following properties [158]:

Table 17.1 Properties of passiv and active electromechanical transducers [158].

PASSIVE TRANSDUCER	ACTIVE TRANSDUCER
reversible coupling between mechanical and electrical or magnetical field values	mechanical values control the power-flow between the inner energy source and the output
no inner energy source required	inner energy source required
signal flow is possible in both directions	signal flow is possible in one direction only
only an input of either electrical or mechanical power generates the corresponding output power	even with missing mechanical power at the input and electric output signal exists

Examples for passive transducers of electromechanical networks are:

- electrostatic transducer,
- piezoelectric transducer,
- magnetic transducer,
- electrodynamic transducer,
- piezomagnetic transduver.

They are either represented by a transformatory or a gyratory coupling.

T.A. Kern (ed.), *Engineering Haptic Devices*, DOI 10.1007/978-3-540-88248-0_17,
© Springer-Verlag Berlin Heidelberg 2009

$$\text{transformatory coupling} \qquad \begin{pmatrix} \mu_L \\ \lambda_L \end{pmatrix} = \begin{pmatrix} X & 0 \\ 0 & 1/X \end{pmatrix} \cdot \begin{pmatrix} \mu_K \\ \lambda_K \end{pmatrix}$$

$$\text{gyratory coulping} \qquad \begin{pmatrix} \mu_L \\ \lambda_L \end{pmatrix} = \begin{pmatrix} 0 & Y \\ 1/Y & 0 \end{pmatrix} \cdot \begin{pmatrix} \mu_K \\ \lambda_K \end{pmatrix}$$

In a transformatory network the flow-values μ and differential values λ are coupled to values of the same class. Gyratory networks couple flow- with differential values and vice versa. The indices L and K mark idle mode resp. short circuit operation. The transducer-coefficients X (transformer) and Y (gyrator) have to be derived from the pyhsical interdependencies. Further information about the identification of transducer-coefficients and a method for the solution of electromechanical networks can be found in [158].

Example:

The basic model of an electrodynamic transducer (fig. 17.1) is identical to a mechanical oscillator. A cylindric disc held by the elasticity n with the mass m and the coil-length l with the turn-number w is able to move along its axis in an air gap. The circular air-gap is filled with the magnetic flux density B_0 of a permanent magnet. A current in the coil generates a LORENTZ-force along the coil's axial direction.

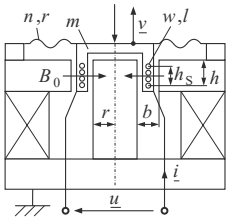


Fig. 17.1 Sketch of an electrodynamic actuator [158]

This LORENTZ-force may be described by the following two equations:

$$\underline{F} = B_0 \cdot l \cdot \underline{i} \qquad \text{und} \qquad \underline{u} = B_0 \cdot l \cdot \underline{v} \tag{17.1}$$

Consequently the electrodynamic transducer is a transformatory coupling of an electrical and a mechanical networt, as the flow-values i and F, such as the differential values v and u are coupled with each other.

The electrodynamic transducer from figure 17.1 is represented by its mechanical equivalent circuit in figure 17.2:

The transducer coefficients are given by the following equations:

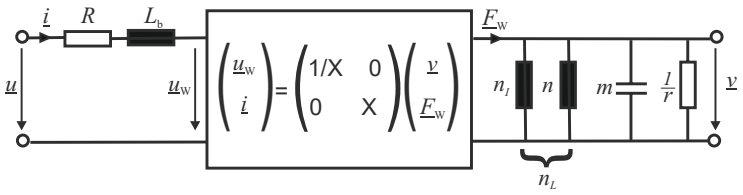


Fig. 17.2 Electromechanical transformer visualized in the electromechanical schematic diagram [158]

$$X = \frac{1}{B_0 \cdot l} \qquad \text{transformatory coupling constant}$$

$$L_b = \frac{\mu_0 \cdot w^2 \cdot 2\pi r \cdot h}{b}\left(1 - \frac{2}{4}\frac{h_s}{h}\right) \qquad \text{inductivity of the coil, in blocked position}$$

$$R = \frac{\rho l}{A}, l = 2\pi r \cdot w \qquad \text{resistance of the coil}$$

$$n_L = n \qquad \text{elasticity}$$

$$m = \rho \cdot V_S \qquad \text{mass of the coil}$$

$$r = \frac{1}{\omega_0 \cdot Q \cdot n_L}, \omega_0^2 = \frac{1}{n_L \cdot m}$$

For transformation of the elements to the other side of the electromechanical network the following transformation laws are given:

- A series circuit stays a series circuit
- A parallel circuit stays a parallel circuit

The transformation from the mechanical to the electrical side gives:

$$C_m = \frac{m}{X^2}$$

$$L_n = \frac{n}{X^2}$$

$$R_r = \frac{1}{r \cdot X^2}$$

$$\underline{Z} = \frac{\underline{h}}{X^2} = \frac{1}{\underline{z} \cdot X^2}$$

The transformation from the electrical to the mechanical side results in:

$$
\begin{aligned}
m_C &= \frac{C}{X^2} \\
n_I &= L \cdot X^2 \\
r_R &= \frac{1}{R \cdot X^2} \\
\underline{h} &= \underline{Z} \cdot X^2 \\
\underline{z} &= \frac{1}{Z \cdot X^2}
\end{aligned}
$$

Chapter 18
Impedance Values of Grasps

The following tables provide the parameter for the model given by figure 4.8 and equation 4.7 in chapter 4. They parametrize the different grasping situations discussed in section 4.2.4.

Table 18.1: Mean values of the mechanical impedance model according to figure 4.8 for different grasping siutations.

Grasp/touch	k_1 [N/m]	m_1 [kg]	d_1 [Ns/m]	d_2 [Ns/m]	k_2 [N/m]	k_3 [N/m]	d_3 [Ns/m]	m_2 [kg]
Power grasps								
Cylinder	412.61	1.577	43.43	33.06	31271	15007	182.77	0.13
Sphere	2500.7	4.32	45.72	31.35	21033	9743	150.60	0.098
Ring 17.71	10.0	0.0032	31.35	5843.7	2906.7	34.54	0.016	
Precision grasps								
Pen 45°	1357.1	1.7376	23.38	3.269	36672	3544.6	12.22	0.029
Pen vertical	44.73	5.44	4.55	17.92	17794	1782.7	12.92	0.029
Pen horizontal	212.49	3.26	7.56	8.15	22092	3672.7	13.73	0.043
Finger								
normal 2ṁm	203.21	75.02	1.0854	3.1672	6656.0	478.73	8.3689	0.0114
normal 15ṁm	0.091	37.28	3.79	3.18	9273.5	839.92	12.22	0.018
shear lateral	54.5	10.0	0.323	4.88	12935	191.62	4.4342	0.0178
shear distal	77.56	9.892	9.443	3.003	22874	2004	4.0377	0.0195
shear 45°	1053.0	90.44	5.47	7.16	26854	1090.2	15.26	0.006

T.A. Kern (ed.), *Engineering Haptic Devices*, DOI 10.1007/978-3-540-88248-0_18,
© Springer-Verlag Berlin Heidelberg 2009

Table 18.2: Linear interpolated dependencies of the model's parameters from grasp- and touch-forces according to figure 4.8 for different grasping situations. Interpolation according to $c = a \cdot F + b$

Grasp/touch		k_1 [N/m]	m_1 [kg]	d_1 [Ns/m]	d_2 [Ns/m]	k_2 [N/m]	k_3 [N/m]	d_3 [Ns/m]	m_2 [kg]
Power grasps									
Cylinder	$a =$	-62.4	-0.216	1.46	-0.409	-365	1330.0	3.27	0.0043
	$b =$	1360	4.88	21.20	39.3	36800	5300.0	133	0.065
Sphere	$a =$	-49.0	-0.111	-.0359	-0.788	13.3	109.0	4.94	.00015
	$b =$	3250	6.01	46.3	43.4	20800	8090.0	75.2	0.096
Ring	$a =$	2.26	$-2.26 10^{-16}$	-.0054	0.143	304.0	150.0	1.72	-.0003
	$b =$	-14.0	10.0	0.107	5.47	1590.0	811.0	10.5	0.206
Precision grasps									
Pen 45°	$a =$	-74.3	0.0616	-0.776	0.0247	-134.0	363.0	0.551	0.00372
	$b =$	1840.0	1.34	28.4	3.11	37500	1190.0	8.64	0.00447
Pen vertical	$a =$	-21.4	-1.26	1.69	0.428	1460.0	204.0	0.759	0.00465
	$b =$	584.0	13.6	6.46	15.1	82.80	454.0	7.99	-.0087
Pen horizontal	$a =$	-27.5	0.56	1.16	-0.619	380.0	229.0	0.409	.00883
	$b =$	391	-0.371	0.193	12.2	19600	2190	11.1	-.0144
Finger									
normal 2mm	$a =$	-124.0	15.2	-0.088	-0.106	-1350	-36.3	1.84	-0.002
	$b =$	606.0	25.5	2.09	3.2	11000	361.0	2.39	0.0180
normal 15mm	$a =$	0.187	-5.95	0.861	-0.233	-1940	374.0	1.48	.000675
	$b =$	0.0311	56.6	0.993	3.94	15600	375	4.75	0.0159
shear lateral	$a =$	-14.0	-106^{-10}	0.177	0.509	1250.0	63.3	0.363	0.00141
	$b =$	100.0	10.0	-0.558	3.23	8860	14.1	3.25	0.0133
shear distal	$a =$	-291.0	0.0571	1.61	-0.711	-8590	367.0	0.266	0.00405
	$b =$	1720	9.71	4.22	5.31	508000	811.0	3.17	0.00636
shear 45°	$a =$	-54.7	-1.94	0.469	1.65	-299.0	74.5	-.0776	.000295
	$b =$	1230.0	96.8	3.95	1.73	27800	848.0	15.5	.00538

References

1. X3D, H-Anim, and VRML 97 Specifications.
   web|3D: http://www.web3d.org/x3d/specifications/.
2. ADAMS, R.; HANNAFORD, B. A. Control law design for haptic interfaces to virtual reality. *Control Systems Technology, IEEE Transactions on 10*, 1 (Jan. 2002), 3 – 13.
3. ADAMS, R. J. *Stable Haptic Interaction with Virtual Environments*. PhD thesis, University of Washington, Washington, 1999.
4. AG, E. R. A. X15G Datasheet. Datenblatt.
5. ALLAM, H. Single element based tactile display. *Mechatronics, 2004. ICM '04. Proceedings of the IEEE International Conference on* (2004), 261 – 267.
6. ATI. *Industrial Automation - Multi Axis Force/Torque Sensors - http://www.ati-ia.com/products/ft/sensors.asp*. 4. Juni 2008.
7. BALLAS, R. G. *Statisches und dynamisches Verhlante piezoelektischer Multilayer-Biegeaktoren und Beitrag zur Sensorintegration*. PhD thesis, Technische Universität Darmstadt, 2006.
8. BALLAS, R. G. *Piezoelectric multilayer beam bending actuators: static and dynamic behavior and aspects of sensor integration*. Microtechnology and MEMS. Springer, Berlin [u.a.], 2007.
9. BAO, M.-H. *Micro mechanical transducers: pressure sensors, accelerometers and gyroscopes*, 2 ed., vol. 8 of *Handbook of sensors and actuators*. Elsevier, Amsterdam [u.a.], 2004.
10. BAR-COHEN, Y. *Electroactive polymer (EAP) actuators as artificial muscles - reality, potential, and challenges*, vol. 98 of *SPIE Press monograph*. SPIE Press, Bellingham, Wash, 2001.
11. BAR-COHEN, Y. *Biomimetics - biologically inspired technologies*. CRC, Taylor und Francis, Boca Raton [u.a.], 2006.
12. BAR-COHEN, Y.; MAVROIDIS, C.; BOUZIT, M.; DOLGIN, B.; HARM, D. L.; KOPCHOK, G. E.; WHITE, R. Virtual reality robotic telesurgery simulations using MEMICA haptic system. In *8th Annual International Symposium on Smart Structures and Material - EAPAD* (2001).
13. BARAFF, D.; WITKIN, A.; KASS, M. Physically based modelling, course notes 36. In *ACM SIGGRAPH '99* (August 1999).
14. BAUMANN, K.; LANZ, H. *Mensch-Maschine-Schnittstelle elektronischer Geräte*. Springer-Verlag, Berlin, 1998.
15. BAUMANN, R. *Haptic Interface for Virtual Reality Based Laparoscopic Surgery Training Environment*. PhD thesis, Département de Microtechnique, Polytechnique Fédérale de Lausanne, 1997.
16. BELFORD, J. D. *The stepped horn - Technical Publication TP-214*. Morgan Electro Ceramics, ...
17. BERKELMAN, P. J.; BUTLER, Z. J.; HOLLIS, R. L. Design of a hemispherical magnetic levitation haptic interface device. In *ASME International Mechanical Engineering Congress and Exposition* (Atlanta, Nov. 1996 1996), ASME, Ed., pp. 483–488.
18. BERTONYMUS. *Wikipedia - Sinn (Wahrnehmung)*. 19.09.2005 2005.
19. BEYELER, F.; MUNTWYLER, S.; NAGY, Z.; GRAETZEL, C.; MOSER, M.; NELSON, B. Design and calibration of a MEMS sensor for measuring the force and torque acting on a magnetic microrobot. *Journal of Micromechanics and Microengineering, IOP 18*, 025004 (Dezember 2007).
20. BICCHI, A. *The sense of touch and its rendering - progress in haptics research*, vol. Vol. 45 of *Springer tracts in advanced robotics*. Springer, Berlin ;Heidelberg, 2008.
21. BICCHI, A.; RAUGI, M.; RIZZO, R.; SGAMBELLURI, N. Analysis and design of an electromagnetic system for the characterization of magneto-rheological fluids for haptic interfaces. *Magnetics, IEEE Transactions on 41*, 5 (2005), 1876 – 1879.

22. BIET, M.; GIRAUD, F.; SEMAIL, B. New tactile devices using piezoelectric actuators. In *ACTUATOR 2006, 10th International Conference on New Actuators* (2006), pp. 989 – 992.
23. BÉKÉSY, G. V. *Physiologie der Sinneshemmung*. Goldmann, München, 1970.
24. BLEICHER, F. *Parallelkinematische Werkzeugmaschinen*. Neuer Wissenschaftlicher Verlag, Graz, 2003.
25. BLUME, H.-J.; BOELCKE, R. *Mechanokutane Sprachvermittlung*, als ms ed., vol. 137 of *Reihe 10*. VDI-Verl., Düsseldorf, 1990.
26. BOTSIS, J.; HUMBERT, L.; COLPO, F.; GIACCARI, P. Embedded fiber Bragg grating sensor for internal strain measurements in polymeric materials. *Optics and Lasers in Engineering 43* (2004).
27. BRAND, U.; BÜTTGENBACH, S. Taktile dimensionelle Messtechnik für Komponenten der Mikrosystemtechnik. *Technisches Messen 12* (2002).
28. BRAY, A.; BARBATO, G.; LEVI, R. *Theory and practice of force measurement*. Monographs in Physical Measurement. Academic Press, London [u.a.], 1990.
29. BRISSAUD, M. Characterization of piezoceramics. *Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on 38*, 6 (Nov. 1991), 603 – 617.
30. BRONSTEIN, I. N.; SEMENDJAJEW, K. A.; MUSIOL, G.; MÜHLIG, H. *Taschenbuch der Mathematik*, 3 ed. Harri Deutsch, Frankfurt am Main, 1997.
31. BROOKS, T. L. Telerobotic Response Requirements. *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics* (1990), 113–120.
32. BÖSE, H.; MONKMANN, G. J.; FREIMUTH, H.; TUNAYAR, A.; KHALED, W.; BAUMANN, M. A new haptic sensor-actuator system based on electrorheological fluids. In *Actuator 2004: 9th International Conference on New Actuators* (2004), HVG Hanseatische Veranst.-GmbH, Div. Messe Bremen, pp. 300 – 303.
33. BÖSE, H.; TRENDLER, A. Smart fluids - properties and benefit for new electromechanical devices. In *AMAS workshop SMART'03* (2003), pp. 329 – 336.
34. BURDEA, G. *Force and touch feedback for virtual reality*. A Wiley-Interscience Publication. Wiley, New York [u.a.], 1996.
35. CALDWELL, D.; LAWTHER, S.; WARDLE, A. Multi-modal cutaneous tactile feedback. In *Intelligent Robots and Systems '96, IROS 96, Proceedings of the 1996 IEEE/RSJ International Conference on* (Osaka, 4. - 8. November 1996), vol. 2, pp. 465 – 472 vol.2.
36. CARLSON, J. D.; STANWAY, R.; JOHNSON, A. R. Electro-rheological and magneto-rheological fluids: a state of the art report. In *Actuator 2004: 9th International Conference on New Actuators* (Bremen, 2004), HVG Hanseatische Veranst.-GmbH, Div. Messe Bremen, pp. 283 – 288.
37. CHAPUIS, D.; CHAPUIS, D.; MICHEL, X. A Haptic Knob with a Hybrid Ultrasonic Motor and Powder Clutch Actuator. In *EuroHaptics Conference, 2007 and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems. World Haptics 2007. Second Joint* (2007), X. Michel R. Gassert, Eds., pp. 200 – 205.
38. CHASE, T.; LUO, R. A thin-film flexible capacitive tactile normal/shear force array sensor. In *Industrial Electronics, Control, and Instrumentation, 1995., Proceedings of the 1995 IEEE IECON 21st International Conference on* (Orlando, FL ISBN: 0-7803-3026-9, 6101995), vol. 2, pp. 1196 – 1201 vol.2.
39. COLGATE, J.; BROWN, J. Factors affecting the Z-Width of a haptic display. *Robotics and Automation, 1994. Proceedings., 1994 IEEE International Conference on* (8-13 May 1994), 3205 – 3210 vol.4.
40. COLGATE, J. E.; GRAFING, P. E.; STANLEY, M. C.; SCHENKEL, G. Implementation of stiff virtual walls in force-reflecting interfaces. In *Proceedings of the IEEE Virtual Realtiy Annual International Symposium(VRAIS)* (Seattle, September 1993), pp. 202–2008.
41. CONRAD, H.; FISHER, M.; SPRECHER, A. F. Characterization of the structure of a model electrorheological fluid employing stereology. In *Proc 2nd Int Conf on ER Fluids* (1990).
42. CRANNY, A.; COTTON, D. P. J. .; CHAPPELL, P. H.; BEEBY, S. P.; WHITE, N. M. Thick-film force, slip and temperature sensors for a prosthetic hand. *Measurement Science and Technology, IOP 16* (März 2005), 931–941.

43. DAI, Z.; GIRAUD, F.; LEMAIRE-SEMAIL, B.; MARTINOT, F. A force feedback device actuated by piezoelectric travelling wave ultrasonic motors. In *ACTUATOR 2006, 10th International Conference on New Actuators* (2006), pp. 600 – 603.
44. DEAN, M. H.; DOUGLAS, R. D. H. *Semiconductor and conventional strain gages*. Academic Press Inc., New York, 1962.
45. DENG, K.; ENIKOV, E.; ZHANG, H. Development of a pulsed electromagnetic microactuator for 3D tactile displays. *Advanced intelligent mechatronics, 2007 ieee/asme international conference on* (2007), 1 – 5.
46. DOERRER, C. *Entwurf eines elektromechanischen Systems für flexibel konfigurierbare Eingabefelder mit haptischer Rückmeldung*, vol. 3 of *Mikro- und Sensortechnik*. Inst. für Elektromechan. Konstruktionen, Darmstadt, 2003.
47. EID, M.; MANSOUR, M.; IGLESIAS, R.; EL SADDIK, A. A Device Independent Haptic Player. In *Virtual Environments, Human-Computer Interfaces and Measurement Systems, 2007. VECIMS 2007. IEEE Symposium on* (Ostuni ISBN: 978-1-4244-0820-7, 2007), pp. 7 – 12.
48. EISNER, E. Complete Solutions of the 'Webster' Horn Equation. *The Journal of the Acoustical Societey of America 41*, 4 (2) (1966), 1126 – 1146.
49. EL WAHED, A. K.; SPROSTON, J. L.; STANWAY, R.; WILLIAMS, E. W. An improved model of ER fluids in squeeze-flow through model updating of the estimated yield stress. *Journal of Sound and Vibration 268*, 3 (2003), 581 – 599.
50. ERICSON, C. *Real-Time Collision Detection (The Morgan Kaufmann Series in Interactive 3-D Technology) (The Morgan Kaufmann Series in Interactive 3D Technology)*. Morgan Kaufmann, December 2004.
51. ERP, J. V.; JANSEN, C.; DOBBINS, T.; VEEN, H. V. Vibrotactile Waypoint Navigation at Sea and in the Air: two Case Studies. 166–173.
52. FALLER, A.; [BEARB.], M. S. *Der Körper des Menschen: Einführung in Bau und Funktion*, 13 ed. Thieme, Stuttgart [u.a.], 1999.
53. FERDINAND, P.; MAGNE, S.; DEWYNTER-MARTY, V.; MARTINEZ, C.; OUGEAULT, S.; BUGAUD, M. Applications of Bragg grating sensors in Europe. In *International Conference on Optical Fiber Sensors OFS* (Wiliamsburg, Virginia, 28. - 31. Oktober 1997).
54. FERNANDEZ, J. M.; PERRIARD, Y. Comparative analysis and modeling of both standing and travelling wave ultrasonic linear motor. *IEEE Ultrasonic Symposium* (2003), 1770 – 1773.
55. FLEISCHER, M.; STEIN, D.; MEIXNER, H. Ultrasonic piezomotor with longitudinally oscillating amplitude-transforming resonator. *Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on 36*, 6 (Nov. 1989), 607 – 613.
56. FÖLLINGER, O. *Nichtlineare Regelungen. 2. Anwendung der Zustandsebene*. R. Oldenbourg Verlag GmbH, 1970.
57. FÖLLINGER, O. *Nichtlineare Regelungen. 3. Ljapunow-Theorie und Popow-Kriterium*. R. Oldenbourg Verlag GmbH, 1970.
58. FÖLLINGER, O. *Nichtlineare Regelungen 1. Grundlagen und Harmonische Balance*. R. Oldenbourg Verlag GmbH, 1991.
59. FLUECKIGER, M.; FLUECKIGER, M.; BULLO, M. fMRI compatible haptic interface actuated with traveling wave ultrasonic motor. In *Industry Applications Conference, 2005. Fourtieth IAS Annual Meeting. Conference Record of the 2005* (2005), M. Bullo D. Chapuis, Eds., vol. 3, pp. 2075 – 2082 Vol. 3.
60. FRANKLIN, G.; POWELL, J.; EMAMI-NAEINI, A. *Feedback Control of Dynamic Systems*. Pearson, 2002.
61. FU, C.-Y.; OLIVER, M. Direct measurement of index finger mechanical impedance at low force. In *Eurohaptics Conference, 2005 and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, 2005. World Haptics 2005. First Joint* (2005), pp. 657 – 659.
62. FUNG, R.-F.; YAO, C.-M.; CHANG, D.-G. Dynamic and contact analysis of a bimodal ultrasonic motor. *Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on 46*, 1 (Jan. 1999), 47 – 60.

63. FUNG, Y.-C. *Biomechanics: mechanical properties of living tissues*, 2 ed. Springer, New York [u.a.], 1993.
64. GEHIN, C., B. C.; TEISSYRE, Y. Design and characterization of new force resonant sensor. *Sensor and Actuators A Physical 84* (2000), 65–69.
65. GERLING, G.; THOMAS, G. The effect of fingertip microstructures on tactile edge perception. In *Haptic Interfaces for Virtual Environment and Teleoperator Systems, 2005. WHC 2005. First Joint Eurohaptics Conference and Symposium on* (2005), pp. 63 – 72.
66. GESCHEIDER, G. A. *Psychophysics: the fundamentals*, 3 ed. Erlbaum, Mahwah, NJ [u.a.], 1997.
67. GESCHEIDER, G. A.; BOLANOWSKI, S. J.; HARDICK, K. R. The frequency selectivity of information-processing channels in the tactile sensory system. *Somatosens Mot Res 18*, 3 (2001), 191–201. Hamilton College, Department of Psychology, Clinton, NY 13323, USA. ggeschei@hamilton.edu.
68. GESCHEIDER, G. A.; BOLANOWSKI, S. J. J.; VERRILLO, R. T.; ARPAJIAN, D. J.; RYAN, T. F. Vibrotactile intensity discrimination measured by three methods. *J Acoust Soc Am 87*, 1 (1990), 330–8. Hamilton College, Clinton, New York 13323.
69. GESCHEIDER, G. A.; SKLAR, B. F.; VAN DOREN, C. L.; VERRILLO, R. T. Vibrotactile forward masking: psychophysical evidence for a triplex theory of cutaneous mechanoreception. *J Acoust Soc Am 78*, 2 (1985), 534–43.
70. GETZINGER, G. *Haptik - Rekonstruktion eines Verlustes*. Profil, Wien, 2006.
71. GHODDSI, R.; BEEBE, D.; WHITE, V.; DENTON, D. Development of a tangential tactor using a LIGA/MEMS linear microactuator technology Microelectromechanical System (MEMS). In *Microelectromechanical System (MEMS), DSC-Vol. 59* (Atlanta, 1996), pp. 379 – 386.
72. GIRAUD, F.; SEMAIL, B. Analysis and phase control of a piezoelectric traveling-wave ultrasonic motor for haptic stick application. *Industry Applications, IEEE Transactions on 40*, 6 (2004), 1541 – 1549.
73. GOETHALS, P.; SETTE, M. M.; REYNAERTS, D.; BRUSSEL, E. V. Flexible Elastoresistive Tactile Sensor for Minimally Invasive Surgery - Eurohaptics Conference. Springer, Berlin/Heidelberg, 2008, pp. 573–579.
74. GOLDSTEIN, E. B. *Sensation and Perception*, 7 ed. Wadsworth Publishing Co., Belmont, California, January 2006.
75. GONG, X.; WU, J.; HUANG, X.; WEN, W.; SHENG, P. Influence of liquid phase on nanoparticle-based giant electrorheological fluid. *Nanotechnology 19*, 16 (2008), 165602 (7pp).
76. GOSLINE, A. H.; CAMPION, G.; HAYWARD, V. On The Use of Eddy Current Brakes as Tunable, Fast Turn-On Viscous Dampers for Haptic Rendering. In *Eurohaptics Conference* (Paris, 2006), pp. 229–234.
77. GRIMALDI, G.; LAMMERTSE, P.; VAN DEN BRABER, N.; MEULEMAN, JOS, M. M. A New Myohaptic Device to Asses Wrist Function in the Lab and in the Clinic - The Wristalyzer - Eurohaptics Conference. In *Haptics: Perception, Devices and Scenarios*, M. Ferre, Ed. Springer, Berlin/Heidelberg, 2008.
78. GRUNWALD, M. *Der bewegte Sinn: Grundlagen und Anwendungen zur haptischen Wahrnehmung*. Birkhäuser, Berlin, 2001.
79. GRUNWALD, M. *Human haptic perception: basics and applications*. Birkhäuser, Berlin, 2008.
80. HAGEDORN, P.; SATTEL, T.; SPEZIARI, D.; SCHMIDT, J.; DIANA, G. The importance of rotor flexibility in ultrasonic traveling wave motors. *Smart Materials and Structures 7* (1998), 352 – 368.
81. HAGOOD, N. W.; MCFARLAND, A. J. Modeling of a piezoelectric rotary ultrasonic motor. *Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on 42*, 2 (March 1995), 210 – 224.
82. HAJIAN, A. Z.; HOWE, R. D. Identification of the Mechanical Impedance at the Human Finger Tip. *Journal of Biomechanical Engineering 119* (Feb. 1997 1997), 109–114.

83. HALE, K.; STANNEY, K. Deriving haptic design guidelines from human physiological, psychophysical, and neurological foundations. *Computer Graphics and Applications, IEEE 24*, 2 (March-April 2004), 33 – 39.
84. HANNAFORD, B.; ANDERSON, R. Experimental and simulation studies of hard contact in force reflecting teleoperation. *Robotics and Automation, 1988. Proceedings., 1988 IEEE International Conference on* (24-29 April 1988), 584 – 589 vol.1.
85. HANNAFORD, B.; RYU, J.-H. Time-domain passivity control of haptic interfaces. *Robotics and Automation, IEEE Transactions on 18*, 1 (Feb. 2002), 1 – 10.
86. HASEGAWA, Y.; SHIKIDA, M.; ET AL. An active tactile sensor for detecting mechanical characteristics of contacted objects. *Journal of Micromechanics and Microengineering, IOP 16* (2006), 1625–1632.
87. HASSER, C. J. Force-Reflecting Anthropomorphic Hand Masters. *Report AL/CF-TR-1995-0110, Armstrong Laboratory, Wright-Patterson Air Force Base, Ohio* (1995).
88. HASSER, C. J.; DANIELS, M. W. Tactile Feedback with Adaptive Controller for a Force-Reflecting Haptic Display - Part I: Design. In *Proceedings of the Fifteenth Southern Biomedical Engineering Conference* (1996), pp. 526–529.
89. HAYWARD, V. A Brief Taxonomy of Tactile Illusions and Demonstrations That Can Be Done In a Hardware Store. *Brain Research Bulletin 756*, 6 (2008), 742–752.
90. HAYWARD, V.; ASTLEY, O. R. Performance Measures for Haptic Interfaces. *Robotics Research: The 7th Int. Symposium 1* (1996), 195–207.
91. HAYWARD, V.; CRUZ-HERNANDEZ, M. Tactile display device using distributed lateral skin stretch. In *Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, IMECE 2000 Conference* (2000), J. Wikander, Ed.
92. HE, S.; CHEN, W.; TAO, X.; CHEN, Z. Standing wave bi-directional linearly moving ultrasonic motor. *Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on 45*, 5 (Sept. 1998), 1133 – 1139.
93. HELIN, P.; SADAUNE, V.; DRUON, C.; TRITSCH, J.-B. *Linear ultrasonic motors using surface acoustic waves mechanical model for energy transfer*, vol. 2 of *Solid State Sensors and Actuators, 1997. TRANSDUCERS '97 Chicago., 1997 International Conference on*. Chicago, IL, 16-19 June 1997.
94. HEYDT, R.; CHHOKAR, S. Refreshable Braille display based on electroactive polymers. In *23rd International Display Research Conference* (2003), pp. 111 – 114.
95. HIRATA, H.; UEHA, S. Design of a traveling wave type ultrasonic motor. *Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on 42*, 2 (March 1995), 225 – 231.
96. HOPP, T. H.; REEVE, C. P. "an algorithm for computing the minimum covering sphere in any dimension".
97. HOPPACH, E.; SCHLAAK, H. F.; WEISSMANTEL, H. Elektrische Kleinantriebe - Skriptum. Tech. rep., Darmstadt, 2006.
98. HOU, L. *Erfassung und Kompensation von Fehlereffekten bei der statischen Kraftmessung mit monolithischen Nd:YAG-Laserkristallen*. PhD thesis, Universität Kassel, 1999.
99. HU, M.; XU, Z.; JIN, L.; WANG, X. *Performance Simulation of Traveling Wave Type Ultrasonic Motor*, vol. 3 of *Electrical Machines and Systems, 2005. ICEMS 2005. Proceedings of the Eighth International Conference on*. 27-29 Sept. 2005.
100. HUANG, X.; SHENG, P.; WEN, W.; YANG, S. Formation of polarized contact layers and the giant electrorheological effect. *International Journal of Modern Physics B (IJMPB) 21*, 28/29 (2007), 4907–4913.
101. HUGONY. Über die Empfindung von Schwingungen mittels des Tastsinnes. *Zeitschrift Biologie 96* (1935), 548–553.
102. HUSTY, M.; KARGER, A.; SACHS, H.; STEINHILPER, W. *Kinematik und Robotik*. Springer, Tokio, 1997.
103. HUSTY, M. L. An algorithm for solving the direct kinematics of general stewart-gough platforms. *Mechanism and Machine Theory 31*, 4 (1996), 365–379.
104. IEEE. *Standard on Piezoelectrity*. 1988.
105. IKEDA, T. *Fundamentals of piezoelectricity*. Oxford Univ. Press, Oxford [u.a.], 1990.

106. IKEI, Y.; IKEI, Y.; SHIRATORI, M. TextureExplorer: a tactile and force display for virtual textures. In *Haptic Interfaces for Virtual Environment and Teleoperator Systems, 2002. HAPTICS 2002. Proceedings. 10th Symposium on* (2002), M. Shiratori, Ed., pp. 327 – 334.
107. IKEI, Y.; YAMADA, M.; FUKUDA, S. Tactile Texture Presentation by Vibratory Pin Arrays Based on Surface Height Maps. In *ASME Dynamic Systems and Control Division*, N. Olgac, Ed., vol. 67. American Society of Mechanical Engineers, New York, NY, 1999, pp. 97 –102.
108. INAUDI, D. SOFO Sensors for Static and Dynamic Measurements. In *1st FIG International Symposium on Engineering Surveys for Construction Works and Structural Engineering* (Nottingham, 28. Juni -1. Juli 2004).
109. ISHII, H. Bottles: A Transparent Interface as a Tribute to Mark Weiser. *IEICE: Transactions on Infomration and Systems E87-D*, 6 (June 2004 2004), 1299–1311299–1311.
110. ISRAR, A.; CHOI, S.; TAN, H. Z. Detection Threshold and Mechanical Impedance of the Hand in a Pen-Hold Posture. In *Intelligent Robots and Systems, 2006 IEEE/RSJ International Conference on* (Beijing, 2006), pp. 472 – 477.
111. ISRAR, A.; CHOI, S.; TAN, H. Z. Mechanical Impedance of the Hand Holding a Spherical Tool at Threshold and Suprathreshold Stimulation Levels. In *EuroHaptics Conference, 2007 and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems. World Haptics 2007. Second Joint* (Tsukaba ISBN: 0-7695-2738-8, 2007), pp. 56 – 60.
112. IWAMOTO, T.; SHINODA, H. Ultrasound tactile display for stress field reproduction - examination of non-vibratory tactile apparent movement. In *Haptic Interfaces for Virtual Environment and Teleoperator Systems, 2005. WHC 2005. First Joint Eurohaptics Conference and Symposium on* (2005), pp. 220 – 228.
113. IWAMOTO, T.; SHINODA, H. Two-dimensional Scanning Tactile Display using Ultrasound Radiation Pressure. In *Haptic Interfaces for Virtual Environment and Teleoperator Systems, 2006 14th Symposium on* (2006), pp. 57 – 61 ISBN: 1–4244–0226–3.
114. JANDURA, L.; SRINIVASAN, M. A. Experiments on Human Performance in Torque Discrimination and Control. *Proceedings of the ASME Winter Annual Meeting, Dynamic Systems and Control, DSC-Vol. 55-1* (1994), 369–375.
115. JANSEN, C.; WENNEMERS, A.; VOS, W.; GROEN, E. FlyTact: A Tactile Display Improves a Helicopter Pilot's Landing Performance in Degraded Visual Environments - Eurohaptics Conference. In *Haptics: Perception, Devices and Scenarios*, M. Ferre, Ed. Springer, Berlin/Heidelberg, 2008, pp. 867–875.
116. JENDRITZA, D. J. *Technischer Einsatz neuer Aktoren: Grundlagen, Werkstoffe, Designregeln und Anwendungsbeispiele*, 2 ed., vol. 484 of *Kontakt Studium*. expert-Verl., Renningen-Malmsheim, 1998.
117. JOLLY, M. R.; CARLSON, J. D. Controllable squeeze film damping using magnetorheological fluid. In *Actuator 96, 5th Int. Conf. on New Actuators* (Bremen, 1996).
118. JONES, L. *Using the Haptic System as a Means of Communication, a Collection of Information - http://bioinstrumentation.mit.edu/jones/pages/Actual/mealani*. 2003.
119. JUNGMANN, M. *Entwicklung elektrostatischer Festkörperaktoren mit elastischen Dielektrika für den Einsatz in taktilen Anzeigefeldern*. PhD thesis, TU Darmstadt, 2004.
120. JUNGMANN, M.; SCHLAAK, H. F. Taktiles Display mit elektrostatischen Polymeraktoren. *Konferenzband des 47. Internationalen Wissenschaftlichen Kolloquiums, Technische Universität Ilmenau* (2002).
121. JUNGNICKEL, U. *Miniaturisierte Positioniersysteme mit mehreren Freiheitsgraden auf der Basis monolithischer Strukturen*. PhD thesis, Darmstadt, 2004.
122. KACZMAREK, K.; WEBSTER, J.; BACH-Y RITA, P.; TOMPKINS, W. Electrotactile and vibrotactile displays for sensory substitution systems. *Biomedical Engineering, IEEE Transactions on 38*, 1 (1991), 1 – 16.
123. KACZMAREK, K. A.; BACH-Y-RITA, P. Tactile Displays (Chapter 9). In *Virtual Environements and Advanced Interface Design*, W. Barfield T. Furness, Eds. Oxford University Press, New-York, 1995, pp. 349–414.

124. KAJIMOTO, H.; KANNO, Y.; TACHI, S. Forehead Electro-tactile Display for Vision Substitution. In *Eurohaptics Conference* (Paris, 2006).
125. KALLENBACH, E.; EICK, R.; QUENDT, P.; STRÖHLA, T.; FEINDT, K.; KALLENBACH, M. *Elektromagnete - Grundlagen, Berechnung, Entwurf und Anwendung*, 3., bearb. u. erg. aufl ed. Vieweg + Teubner, Wiesbaden, 2008.
126. KANT, I. *Anthropologie in pragmatischer Hinsicht*. Reclam, Stuttgart, 1983.
127. KENALEY, G. L.; CUTKOSKY, M. R. Electrorheological fluid-based robotic fingers with tactile sensing. In *Robotics and Automation, 1989. Proceedings., 1989 IEEE International Conference on* (1989), pp. 132 – 136 vol.1.
128. KERDOK, A. E. *Characterizing the Nonlinear Mechanical Response of Liver to Surgical Manipulation*. PhD thesis, Harvard University, Cambridge, May 2006.
129. KERN, T. A. *Haptisches Assistenzsystem für diagnostische und therapeutische Katheterisierungen: "Mit dem Blindenstock durch Adern tasten!"*. PhD thesis, 2006.
130. KERN, T. A.; KLAGES, S.; MEISS, T.; WERTHSCHÜTZKY, R. Study of the Influence of Varying Diameter and Grasp-Forces on Mechanical Impedance for the Grasp of Cylindrical Objects. *Eurohaptics Conference 1* (3.- 6. Juni 2006 2006), 113–118.
131. KERN, T. A.; RÖRUP, H.; WERTHSCHUTZKY, R.; TAMMER, R. A remotely controlled lightweight MRI compatible ultrasonic actuator for micrometer positioning of electrodes during neuroethological primate research / Telemetrisch gesteuerter, MRT-kompatibler Ultraschallaktor zur Mikrometerpositionierung von Elektrode. *Biomed Tech (Berl) 53*, 6 (2008), 292–9. 1Institute for Electromechanical Design, Darmstadt University of Technology, Darmstadt, Germany.
132. KERN, T. A.; SCHAEFFER, A.; WERTHSCHÜTZKY, R. An Interaction Model for the Quantification of Haptic Impressions. In *Haptics: Perception, Devices and Scenarios*, M. Ferre, Ed. Sppringer, Berlin Heidelberg, 2008, pp. 139–145.
133. KERSEY, A. D.; DAVIS, M. A.; PATRICK, H. J., E. A. Fiber Grating Sensors. *Journal of Lightwave Technology 15*, 8 (1997).
134. KG, P. I. P. G. . C. Piezo linear driving mechanism for converting electrical energy into motion, has a group of stacking actuators for driving a rotor in a guiding mechanism, 2003.
135. KHALIL, H. K. *Nonlinear Systems*. Prentice Hall, 2002.
136. KIKUUWE, R.; SANO, A.; MOCHIYAMA, H.; TAKESUE, N.; TSUNEKAWA, K.; SUZUKI, S.; FUJIMOTO, H. The tactile contact lens. In *Sensors, 2004. Proceedings of IEEE* (2004), pp. 535 – 538 vol.2.
137. KIM, K. J.; TADOKORO, S. *Electroactive Polymers for Robotic Applications - Artificial Muscles and Sensors*. Springer eBook Collection, Engineering [Dig. Serial], Springer-11647 [Dig. Serial]. Springer-Verlag London Limited, London, 2007.
138. KIRCHNER, J. Mehrkriterielle optimierung von parallelkinematiken. *Verlag Wissenschaftliche Scripten, Zwickau* (2001).
139. KLAGES, S. *Neue Sensorkonzepte zur Zungendruckmessung*. Diplomarbeit, Technische Universität Darmstadt, Darmstadt, August 2004.
140. KLOSE, S. *Abschätzung der Miniaturisierbarkeit eines Ultraschallaktors*. PhD thesis, Technische Universität Darmstadt, 2006.
141. KONTARINIS, D. A.; HOWE, R. D. Tactile Display of Contact Shape in Dextrous Telemanipulation. *ASME Winter Annual Meeting: Advances in Robotics, Mechatronics and Haptic Interfaces, New Orleans, DSC-Vol. 49* (1993), 81–88.
142. KONYO, M.; TADOKORO, S.; YOSHIDA, A.; SAIWAKI, N. A tactile synthesis method using multiple frequency vibrations for representing virtual touch. In *Intelligent Robots and Systems, 2005. (IROS 2005). 2005 IEEE/RSJ International Conference on* (2005), pp. 3965 – 3971.
143. KORNBLUH, R.; PELRINE, R. Electrostrictive polymer artificial muscle actuators. In *Robotics and Automation, 1998. Proceedings. 1998 IEEE International Conference on* (1998), vol. 3, pp. 2147 – 2154 vol.3.
144. KORNBLUH, R.; PELRINE, R.; JOSEPH, J.; HEYDT, R.; PEI, Q.; CHIBA, S. High-Field electrostriction of elastomeric polymer dielectrics for actuation. In *6th Annual International Symposium on Smart Structures and Material - EAPAD* (1999), pp. 149 – 161.

145. KOYAMA, T.; KOYAMA, T.; TAKEMURA, K. Development of an ultrasonic clutch for multi-fingered exoskeleton haptic device using passive force feedback for dexterous teleoperation. In *Intelligent Robots and Systems, 2003. (IROS 2003). Proceedings. 2003 IEEE/RSJ International Conference on* (2003), K. Takemura T. Maeno, Eds., vol. 3, pp. 2229 – 2234 vol.3.
146. KUCHENBECKER, K.; FIENE, J.; NIEMEYER, G. *Event-based haptics and acceleration matching: portraying and assessing the realism of contact*. 18-20 March 2005.
147. KUNSTMANN, C. *Handhabungssystem mit optimierter Mensch-Maschine-Schnittstelle für die Mikromontage*, als ms ed., vol. 751 of *Reihe 8, Mess-, Steuerungs- und Regelungstechnik*. VDI-Verl., Düsseldorf, 1999.
148. KUSCHEL, M.; KREMER, P.; HIRCHE, S.; BUSS, M. Lossy data reduction methods for haptic telepresence systems. *Robotics and Automation, 2006. ICRA 2006. Proceedings 2006 IEEE International Conference on* (2006), 2933 – 2938.
149. KYUNG, K.-U.; AHN, M.; KWON, D.-S.; SRINIVASAN, M. A compact broadband tactile display and its effectiveness in the display of tactile form. In *Haptic Interfaces for Virtual Environment and Teleoperator Systems, 2005. WHC 2005. First Joint Eurohaptics Conference and Symposium on* (2005), pp. 600 – 601.
150. KYUNG, K.-U.; KYUNG, K.-U.; LEE, J. Y. Haptic Stylus and Empirical Studies on Braille, Button, and Texture Display. *Journal of Biomedicine and Biotechnology vol. 2008*, Article ID 369651 (2007), 11.
151. KYUNG, K.-U.; KYUNG, K.-U.; LEE, J. Y. Pen-like Haptic Interface and Its Application on Touch Screen. In *Robot and Human interactive Communication, 2007. RO-MAN 2007. The 16th IEEE International Symposium on* (2007), J. Y. Lee J.-S. Park, Eds., pp. 9 – 13.
152. KYUNG, K.-U.; KYUNG, K.-U.; PARK, J.-S. Ubi-Pen: Development of a Compact Tactile Display Module and Its Application to a Haptic Stylus. In *EuroHaptics Conference, 2007 and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems. World Haptics 2007. Second Joint* (2007), J.-S. Park, Ed., pp. 109 – 114.
153. LANDAU, K. *Die Arbeit im Dienstleistungsbetrieb: Grundzüge einer Arbeitswissenschaft der personenbezogenen Dienstleistung*. Ulmer, Stuttgart, 1996.
154. LANGETEPE, E.; ZACHMANN, G. *Geometric Data Structures for Computer Graphics*, 1 ed. AK Peters, 2005.
155. LASERCOMPONENTS. *Laser Components Group - Faseroptische Sensoren - http://www.lasercomponents.com/de/1134.html*. 4. Juni 2008.
156. LAWRENCE, D. Stability and transparency in bilateral teleoperation. *Robotics and Automation, IEEE Transactions on 9*, 5 (Oct. 1993), 624 – 637.
157. LAWRENCE, D.; PAO, L.; APHANUPHONG, S. Bow spring/tendon actuation for low cost haptic interfaces. *Haptic Interfaces for Virtual Environment and Teleoperator Systems, 2005. WHC 2005. First Joint Eurohaptics Conference and Symposium on* (18-20 March 2005), 157 – 166.
158. LENK, A.; PFEIFER, G.; WERTHSCHÜTZKY, R. *Elektromechanische Systeme: mechanische und akustische Netzwerke, deren Wechselwirkungen und Anwendungen*. Springer, Berlin [u.a.], 2001.
159. LEONDES, C. T. *MEMS/NEMS - handbook techniques and applications*. Springer, New York, NY, 2006.
160. LEUSCHKE, R.; KURIHARA, E.; DOSHER, J.; HANNAFORD, B. High fidelity multi finger haptic display. *Eurohaptics Conference, 2005 and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, 2005. World Haptics 2005. First Joint* (18202005), 606 – 608.
161. LEVESQUE, V.; LEVESQUE, V.; PASQUERO, J. Braille Display by Lateral Skin Deformation with the STReSS2 Tactile Transducer. In *EuroHaptics Conference, 2007 and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems. World Haptics 2007. Second Joint* (2007), J. Pasquero V. Hayward, Eds., pp. 115 – 120.
162. LI, W. H.; DU, H.; GUO, N. Q.; KOSASIH, P. B. Magnetorheological fluids based haptic device. *Sensor Review 24*, 1 (2004), 68 – 73.

163. LIN, M. C.; CANNY, J. F. Efficient algorithms for incremental distance computation. In *IEEE Conference on Robotics and Automation* (1991), pp. 1008–1014.
164. LUNZE, J. *Regelungstechnik 2*. Springer, Berlin, 2005.
165. LUNZE, J. *Regelungstechnik 1*. Springer, Berlin, 2006.
166. LUO, F.; LIU, J.; MA, N.; MORSE, T. A fiber optic microbend sensor for distributed sensing application in the structural strain monitoring. *Sensors and Actuators A Physical 75* (1999), 41–44.
167. LUTZ, H.; WENDT, W. *Taschenbuch der Regelungstechnik*, 5 ed. Deutsch, Frankfurt am Main, 2003.
168. MAGNUSSEN, BJÖRN; HAGEMANN, B. S. D. D. E. W. K. V. P. Piezomotor and methods for the production and exitement thereof, 30. Okt.. 2003.
169. MCNEELY, W. A.; PUTERBAUGH, K. D.; TROY, J. J. Six degree-of-freedom haptic rendering using voxel sampling. In *SIGGRAPH* (1999), pp. 401–408.
170. MEHLING, J.; COLGATE, J.; PESHKIN, M. Increasing the impedance range of a haptic display by adding electrical damping. In *Eurohaptics Conference, 2005 and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, 2005. World Haptics 2005. First Joint* (2005), pp. 257 – 262.
171. MEISS, T.; KERN, T. A.; KLAGES, S.; WERTHSCHUETZKY, R. Fertigung eines Miniaturkraftsensors mit asymmetrischem Grundkörper zur Anwendung bei Katheterisierungen. In *MikroSystemTechnik Kongress* (Dresden, 2007), VDI-VDE, Ed.
172. MEISSNER, P. *Seminar zu speziellen Themen der optischen Nachrichtentechnik - Simulatorische und experimentelle Untersuchungen optischer WDM Übertragungssysteme*. Skriptum, Technische Universität Darmstadt, Juni 2006.
173. MEISSNER, P. *Optische Nachrichtentechnik I - Skriptum zur Vorlesung*. Skriptum, Technische Universität Darmstadt, 2007.
174. MERLET, J. P. *Parallel Robots*, 2. ed ed. Solid mechanics and its applications. Springer, Dordrecht, 2006.
175. MESCHEDE, D. *Optics, light and lasers: the practical approach to modern aspects of photonics and laser physics*, 2 ed. Optik, Licht und Laser engl. Wiley-VCH-Verl., Weinheim, 2007.
176. MILLER, B.; COLGATE, J.; FREEMAN, R. Environment delay in haptic systems. *Robotics and Automation, 2000. Proceedings. ICRA '00. IEEE International Conference on 3* (24-28 April 2000), 2434 – 2439 vol.3.
177. MILNER, T. E.; FRANKLIN, D. W. Characterization of Multijoint Finger Stiffness: Dependence on Finger Posture and Force Direction. *IEEE Transactions on Biomedical Engineering, Vol. 43, Nr. 11* (1998), 1363–1375.
178. MITRA, P.; GENTRY, D.; NIEMEYER, G. User Perception and Preference in Model Mediated Telemanipulation. In *EuroHaptics Conference, 2007 and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems. World Haptics 2007. Second Joint* (Tsukaba, 2007), pp. 268 – 273.
179. MÜLLER, G.; GROTH, C. *FEM für Praktiker - Grundlagen: Basiswissen und Arbeitsbeispiele zur Finite-Element-Methode mit dem FE-Programm ANSYS Rev. 5.5 : mit CD-ROM, Installationsanleitung und zahlreichen Beispielen von: FEM für Praktiker*, 7 ed., vol. 23 of *Edition expertsoft*. expert-Verl., Renningen-Malmsheim, 2002.
180. MOAL, P. L.; JOSEPH, E.; FERNIOT, J.-C. Mechanical energy transductions in standing wave ultrasonic motors: Analytical modelling and experimental investigations. *European Journal of Mechanics - A/Solids 19*, 5 (2000), 849–871.
181. MURAYAMA, J.; BOURGILA, L.; LUO, Y.; AKAHANE, K.; HASEGAWA, S.; HIRSBRUNNER, B.; SATO, M. SPIDAR G&G: A Two-Handed Haptic Interface for Bimanual VR Interaction. In *Eurohaptics*, vol. 1. Universität München, München, 2004, pp. 138–146.
182. NAKASHIGE, M.; HIROTA, K.; HIROSE, M. Linear actuator for high-resolution tactile display. *Robot and Human Interactive Communication, 2004. ROMAN 2004. 13th IEEE International Workshop on* (2004), 587 – 590.

183. NAKATANI, M.; HOWE, R. D.; TACHI, S. The Fishbone Tactile Illusion. In *Eurohaptics Conference* (Paris, 3.- 6. Juni 2006 2006), Academie de Technologies.
184. NAKATANI, M.; HOWE, R. D.; TACJI, S. The Fishbone Tactile Illusion. In *Eurohaptics* (Paris, 2006), IEEE, Ed.
185. NEUGEBAUER, R. *Parallelkinematische Maschinen: Entwurf, Konstruktion, Anwendung*. VDI-Buch. Springer, Berlin [u.a.], 2006.
186. NÉRET, G. *Erotik in der Kunst des 20. Jahrhunderts*. Taschen, Tokyo, 1998.
187. NÜHRMANN, D. *Das Groß e Werkbuch Elektronik*, 7. ed. Franzis' Verlag, Poing, 1998.
188. OGUZTORELI, M. N.; STEIN, R. B. Optimal task performance of antagonistic muscles. *Biol Cybern 64*, 2 (1990), 87–94. Department of Mathematics, University of Alberta, Edmonton, Canada.
189. PAHL, G.; BEITZ, W.; FELDHUSEN, J. *Konstruktionslehre: Grundlagen erfolgreicher Produktentwicklung; Methoden und Anwendung*, 7 ed. Springer, Berlin [u.a.], 2007.
190. PALAIS, J. C. *Fiber optic communications*, 5 ed. Pearson Prentice-Hall, Upper Saddle River, NJ, 2005.
191. PANDEY, N.; YADAV, B. C. Embedded fibre optic microbend sensor for measurement of high pressure and crack detection. *Sensors and Actuators A Physical 128* (2006), 33–36.
192. PANG, X. D.; TAN, H. Z.; DURLACH, N. I. Manual Discrimination of Force Using Active Finger Motion. *Perception Psychophysics, Vol. 49, Nr. 6* (1991), 531–540.
193. PARE, M.; ELDE, R.; MAZURKIEWICZ, J. E.; SMITH, A. M.; RICE, F. L. The Meissner corpuscle revised: a multiafferented mechanoreceptor with nociceptor immunochemical properties. *J Neurosci 21*, 18 (2001), 7236–46. Departement de Physiologie, Universite de Montreal, Montreal, Quebec, Canada H3C 3J7.
194. PARTHASARATHY, M.; KLINGENBERG, D. J. Electrorheology: Mechanisms and models. *Materials Science and Engineering: R: Reports 17*, 2 (1996), 57 – 103.
195. PARTSCH, U. *LTCC-kompatible Sensorschichten und deren Applikation in LTCC-Drucksensoren*, 1 ed., vol. 9 of *Elektronik-Technologie in Forschung und Praxis*. Detert, Templin/Uckermark, 2002.
196. PEI, Q.; ROSENTHAL, M. A.; PELRINE, R.; STANFORD, S.; KORNBLUH, R. Multifunctional electroelastomer roll actuators and their application for biomimetic walking robots. In *Proceedings of SPIE, 5051 [31]* (2003), pp. 281 – 290.
197. PEIRS, J.; CLIJNEN, J.; REYNAERTS, D.; VAN BRUSSEL, H.; HERIJGERS, P.; CORTEVILLE, B.; BOONE, S. A micro optical force sensor for force feedback during minimally invasive robotic surgery. *Sensors and Actuators A: Physical 115*, 2-3 (2004), 447–455.
198. PELRINE, R.; KORNBLUH, R. Electromechanical transduction effects in dielectric elastomers: actuation, sensing, stiffness modulation and electric energy generation. In *Dielectric Elastomers as Electromechanical Transducers*, F. Carpi, D. De Rossi, R. Kornbluh, P. Sommer-Larsen, Eds., 1 ed. Elsevier, Amsterdam, 2008, p. 344.
199. PHILLIPS, R. W. *Engineering applications of fluids with a variable yield stress*. PhD thesis, University of California, Berkeley, 1969.
200. PING, L.; YUMEI, W. An arbitrarily distributed tactile sensor array using piezoelectric resonator. In *Instrumentation and Measurement Technology Conference, 1996. IMTC-96. Conference Proceedings. 'Quality Measurements: The Indispensable Bridge between Theory and Reality'., IEEE* (Brussels ISBN: 0-7803-3312-8, 1996), vol. 1, pp. 502 – 505 vol.1.
201. PUERS, R. Capacitive sensors: when and how to use them. *Sensors and Actuators A: Physical 37-38* (1993), 93–105.
202. RAJAGOPAL, K. R.; RUZICKA, M. Mathematical modeling of electrorheological materials. *Continuum Mechanics and Thermodynamics 13*, 1 (2001), 59 – 78.
203. RASKIN, J. *The Humane Interface*. Addison-Wesley, Reading, Massachusetts, 2000.
204. RAUSCH, J. *Analyse der mechanischen Eigenschaften von Lebergewebe bei intrakorporaler Interaktion*. Diplomarbeit, Technische Universität Darmstadt, Darmstadt, 2005.
205. REISSELL, L.-M.; PAI, D. K. High Resolution Analysis of Impact Sounds and Forces. In *EuroHaptics Conference, 2007 and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems. World Haptics 2007. Second Joint* (Tsukaba, 2007), pp. 255 – 260.

206. Ren, K.; Liu, S.; Lin, M.; Wang, Y.; Zhang, Q. M. A compact electroactive polymer actuator suitable for refreshable Braille display. *Sensors and Actuators A: Physical 143*, 2 (2008), 335 – 342.
207. Rey, P.; Charvet, P.; et al. A high density capacitive pressure sensor array for fingerprint sensor application. In *International Conference on Solid-State Sensors and Actuators* (Chicago, 1997), IEEE, Ed.
208. Riener, R.; Villgrattner, T.; Kleiser, R.; Nef, T.; Kollias, S. fMRI-Compatible Electromagnetic Haptic Interface. In *Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the* (Shanghai ISBN: 0-7803-8741-4, 2005), pp. 7024 – 7027.
209. Röse, A.; Kern, T. A.; Eicher, D.; Schemmer, B.; Schlaak, H. F. INKOMAN An intracorporal manipulator for minimally invasive surgery. In *Proceedings Biomedizinische Technik* (Zürich, 06.-09.09. 2006).
210. Rosen, J.; Solazzo, M.; Hannaford, M. D. B. Task Decomposition of Laparoscopic Surgery for Objective Evaluation of Surgical Residents Learning Curve Using Hidden Markov Model. *Computer Aided Surgery 7* (2002).
211. Ruspini, D. C.; Kolarov, K.; Khatib, O. The haptic display of complex graphical environments. In *SIGGRAPH '97: Proceedings of the 24th annual conference on Computer graphics and interactive techniques* (New York, NY, USA, 1997), ACM Press/Addison-Wesley Publishing Co., pp. 345–352.
212. Russell, R. A. A tactile sensory skin for measuring surface contours. In *Tencon 1992 Region 10 Conference* (Melbourne, November 1992), IEEE, Ed.
213. Saraf, R.; Maheshwari, V. High-Resolution Thin-Film Device to Sense Texture by Touch. Tech. Rep. 5779, 9. Juni 2006.
214. Sattel, T. Dynamics of ultrasonic motors. IV, 167.
215. Schaumburg, H. *Sensoren*, vol. 3 of *Werkstoffe und Bauelemente der Elektrotechnik*. Teubner, Stuttgart, 1992.
216. Schimkat, J.; Kiesewetter, L.; Gevatter, H.; Arndt, F.; Steckenborn, A.; Schlaak, H. F. Moving Wedge Actuator: An Electrostatic Actuator for Use in a Microrelay. In *MICRO SYSTEM Technologies '94, 4th International Conference and Exhibition on Micro, Electro, Opto, Mechanical Systems and Components* (1994), VDE-Verlag, pp. 989 – 996.
217. Schimmel, P.; Noever, P. *Out of actions: zwischen Performance und Objekt 1949 - 1979*. Cantz, Ostfildern, 1998. Kämpfen mit Schlamm.
218. Schlaak, H.; Hoppach, E.; Weissmantel, H.; Eicher, D. a. *Elektrische Kleinantriebe - Skript*. Vorlesungsmaterialien. Eigenverlag, Darmstadt, 2007.
219. Schmidt, R. F. *Neuro- und Sinnesphysiologie*, 4 ed. Springer, Berlin [u.a.], 2000.
220. Schnell, W.; Gross, D.; Hauger, W. *Technische Mechanik 2: Elastostatik*, 7 ed. Springer, Berlin [u.a.], 2002.
221. Schreier-Alt, T. *Polymerverkapselung mechatronischer Systeme: Charakterisierung durch eingebettete Faser-Bragg-Gitter-Sensoren*. PhD thesis, TU Berlin, 2006.
222. Sektion für Minimal Invasive Chirurgie Tübingen, U. T. Verbundprojekt TAMIC - Entwicklung eines taktilen Mikrosensors für die Minimal Invasive Chirurgie - Schlussbericht. Tech. rep., 1999.
223. Sergio, M.; Manaresi, N.; Campi, F.; et al. A dynamically reconfigurable monolithic CMOS pressure sensor for smart fabric. *IEEE Journal of solid-state circuits 38*, 6 (2003), 966–968.
224. Serina, Elaine, R.; Mote, C.; Rempel, D. Force Response of the Fingertip Pulp to Repeated Compression - Effects of Loading Rate, Loading Angle and Anthropometry. *Journal of Biomechanics 30*, 10 (1997 1997), 1035–1040.
225. Sessler, G. M. Akkustik I - Skriptum zur Vorlesung. 2000.
226. Shahabi, C.; Ortega, A.; Kolahdouzan, M. R. A comparison of different haptic compression techniques. *Multimedia and Expo, 2002. ICME'02. Proceedings. 2002 IEEE International Conference on 1* (2002).

227. SHENG, P. Mechanism of the giant electrorheological effect. *International Journal of Modern Physics B (IJMPB) 19*, 7/9 (2005), 1157–1162.
228. SHERRIT, S.; DOLGIN, B. P.; BAR-COHEN, Y.; PAL, D.; KROH, J.; PETERSON, T. Modeling of horns for sonic/ultrasonic applications. *Ultrasonics Symposium, 1999. Proceedings. 1999 IEEE 1* (1999), 647 – 651 vol.1.
229. SHIMOGA, K. B. A Survey of Perceptual Feedback Issues in Dextrous Telemanipulation: Part I. Finger Force Feedback. *Proceedings of the IEEE Virtual Reality Annual International Symposium (VRAIS), Seattle, Washington* (1993), 263–270.
230. SHIMOGA, K. B. A Survey of Percepual Feedback Issues in Dextrous Telemanipulation: Part II. Finger Touch Feedback. *Proceedings of the IEEE Virtual Reality Annual International Symposium (VRAIS), Seattle, Washington* (1993), 271–279.
231. SHUAI, L.-G.; KUANG, Y.-H.; ZHOU, Z.-T. Research on robot tactile display based on water jet technology. In *Robotics and Biomimetics (ROBIO). 2005 IEEE International Conference on* (Shatin ISBN: 0-7803-9315-5, 2005), pp. 676 – 680.
232. SIEGEL, M. Tactile display development: the driving-force for tactile sensor development. *Haptic Virtual Environments and Their Applications, IEEE International Workshop 2002 HAVE* (2002), 115 – 118.
233. SIMONE, C. *Modelling of neddle insertion forces of percutaneous therapies*. Diplomarbeit, Johns Hopkins University, Baltimore, 2002.
234. SLOTINE, J.-J. E.; LI, W. *Applied Nonlinear Control*. Prentice Hall, 1991.
235. SMI. *Silicon Microstructures Incooperated - Pressure Sensors Products - http://www.si-micro.com/pressure-sensor-products.html*. 4. Juli 2008.
236. SMITH, C. U. *The biology of sensory systems*. Wiley, Chichester [u.a.], 2000.
237. SMITH, C., S. Piezoresistance Effect in Germanium and Silicon. *Phys. Rev., American Physical Society 94*, 1 (April 1954), 42–49.
238. Newsclae Technologies - Ingenioulsy small motion systems.
239. SRINIVASAN, M. A.; CHEN, J.-S. Human Performance in Controlling Normal Forces of Contact with Rigid Objects. *ASME Winter Annual Meeting: Advances in Robotics, Mechatronics and Haptic Interfaces, New Orleans, DSC-Vol. 49* (1993), 119–125.
240. STAVROULIS, S. *Rechnergestützter Entwurf von piezoresistiven Silizium-Drucksensoren mit realem mechanischem Wandler*. PhD thesis, 2004.
241. STELZMANN, U.; GROTH, C.; MÜLLER, G. *FEM für Praktiker - Strukturdynamik: Basiswissen und Arbeitsbeispiele zu FEM-Anwendungen der Strukturdynamik von: FEM für Praktiker*, 3 ed., vol. 44 of *Edition expertsoft*. expert-Verl., Renningen, 2002.
242. STEVENS, S. S. *Psychophysics: introduction to its perceptual, neural, and social prospects*. A Wiley-Interscience publication. Wiley, New York [u.a.], 1975.
243. SU, L.; CHIANG, K. S.; LU, C. Microbend-induced mode coupling in a graded-index multimode fiber. *APPLIED OPTICS 44*, 34 (Dezember 2005).
244. SUMMERS, I. R.; CHANTER, C. M. A broadband tactile array on the fingertip. *The Journal of the Acoustical Society of America 112*, 5 (2002/11/00/ 2002), 2118 – 2126.
245. SUN, Y.; NELSON, B.; POTASEK, D.; ENIKOV, E. A bulkmicrofabricated multi-axis capacitive cellular force sensor using transverse comb drives. *Journal of Micromechanics and Microengineering, IOP 12* (2002), 832–840.
246. SUTHERLAND, I. E. The Ultimate Display. In *Proceedings of IFIP Congress* (1965), pp. 506 – 508.
247. TAKAMATSU, R.; TANIGUCHI, T.; SATO, M. Space browsing interface based on head posture information. In *Computer Human Interaction, 1998. Proceedings. 3rd Asia Pacific* (Shonan Village Center ISBN: 0-8186-8347-3, 1998), pp. 298 – 303.
248. TAKASAKI, M.; KURIBAYASHI KUROSAWA, M.; HIGUCHI, T. Optimum contact conditions for miniaturized surface acoustic wave linear motor. *Ultrasonics 38*, 1-8 (2000/3 2000), 51 – 53.
249. TALBOT, W. H.; DARIAN-SMITH, I.; KORNHUBER, H. H.; MOUNTCASTLE, V. B. The sense of flutter-vibration: comparison of the human capacity with response patterns of mechanoreceptive afferents from the monkey hand. *J Neurophysiol 31*, 2 (1968), 301–34.

250. TAMMER, R.; EHRENREICH, L.; JURGENS, U. Telemetrically recorded neuronal activity in the inferior colliculus and bordering tegmentum during vocal communication in squirrel monkeys (Saimiri sciureus). *Behav Brain Res 151*, 1-2 (2004), 331–6. German Primate Centre, Kellnerweg 4, 37077 Gottingen, Germany.
251. TAN, H. Z.; PANG, X. D.; DURLACH, N. I. Manual Resolution of Length, Force, and Compliance. *ASME Advances in Robotics, DSC-Vol. 42* (1992), 13–18.
252. TAN, H. Z.; SRINIVASAN, M. A.; EBERMANN, B.; CHENG, B. Human Factors for the Design of Force-Reflecting Haptic Interfaces. *Proceedings of the ASME Winter Annual Meeting: Dynamic Systems and Control, DSC-Vol. 55-1* (1994), 353–359.
253. TANG, H.; BEEBE, D. J. A microfabricated electrostatic haptic display for persons with visual impairments. *Rehabilitation Engineering, IEEE Transactions on Neural Systems and Rehabilitation 6*, 3 (1998), 241 – 248.
254. TAYLOR, P. M.; POLLET, D. M.; HOSSEINI-SIANAKI, A.; VARLEY, C. J. Advances in an electrorheological fluid based tactile array. *Displays 18*, 3 (1998), 135 – 141.
255. TIETZE, U.; SCHENK, C. *Halbleiter-Schaltungstechnik: / U. Tietze*, 12 ed. Springer, Berlin [u.a.], 2002.
256. TSAGARAKIS, N.; HORNE, T.; CALDWELL, D. SLIP AESTHEASIS: a portable 2D slip/skin stretch display for the fingertip. In *Haptic Interfaces for Virtual Environment and Teleoperator Systems, 2005. WHC 2005. First Joint Eurohaptics Conference and Symposium on* (2005), pp. 214 – 219.
257. TSAI, L.-W. *Robot analysis - the mechanics of serial and parallel manipulators*. A Wiley-Interscience publication. Wiley, New York [u.a.], 1999.
258. TSAI, L. W.; JOSHI, S. Kinematics and optimization of a spatial 3-upu parallel manipulator. *Journal of Mechanical Design 122* (2000), 439.
259. TSAI, L. W.; WALSH, G. C.; STAMPER, R. E. Kinematics of a novel three dof translational platform. *Robotics and Automation, 1996. Proceedings., 1996 IEEE International Conference on 4* (1996).
260. UCHINO, K. *Piezoelectric actuators and ultrasonic motors*. Kluwer, Boston [u.a.], 1997.
261. UCHINO, K.; GINIEWICZ, J. R. *Micromechatronics*. No. 22 in Materials engineering. Dekker, New York [u.a], 2003.
262. UHEA, S. *Recent development of ultrasonic actuators*, vol. 1 of *Ultrasonics Symposium, 2001 IEEE*. Atlanta, GA, 7-10 Oct. 2001.
263. UHEA, S.; TOMIKAWA, Y.; KUROSAWA, M.; NAKAMURA, N. *Ultrasonic motors - Theory and application*. Oxford University Press, 1993.
264. UNBEHAUEN, H. *Regelungstechnik I*. Vieweg und Teubner, 2007.
265. UNBEHAUEN, H. *Regelungstechnik II: Zustandsregelungen, digitale und nichtlineare Regelsysteme*. Vieweg und Teubner, 2007.
266. VALDASTRI, P.; ROCCELLA, S.; BECCAI, L.; CATTIN, E.; MENCIASSI, A.; CARROZZA, M.; DARIO, P. Characterization of a novel hybrid silicon three-axial force sensor. *Sensors and Actuators A Physical 123-124* (2005), 249–257.
267. VAN ERP, J. Vibrotactile spatial acuity on the torso: effects of location and timing parameters. *Haptic Interfaces for Virtual Environment and Teleoperator Systems, 2005. WHC 2005. First Joint Eurohaptics Conference and Symposium on* (2005), 80 – 85.
268. VARADI, P. Piezomotor with a guide, 3. April 2003.
269. VASARHELYI, G.; ADAMA, M.; ET AL. Effects of the elastic cover on tactile sensor arrays. *Sensors and Actuators A Physical 132* (2006), 245–251.
270. VASARHELYI, G.; FODOR, B.; ROSKA, T. Tactile sensing-processing - Interface-cover geometry and the inverse-elastic problem. *Sensors and Actuators A Physical 140* (2007), 8–18.
271. VAZSONYI, E.; ADAM, M.; DUSCO, C.; VIZVARY, Z.; TOTH, A. L.; BARSONY, I. Three-dimensional force sensor by novel alkaline etching technique. *Sensors and Actuators A Physical 123-124* (2005), 620–626.
272. VIDAL VERDÚ, F.; NAVAS GONZALES, R. Thermopneumatic Approach for Tactile Displays. In *Mechatronics & Robotics* (Aachen, Germany, 2004), pp. 394–399.

273. VITRANI, M. A.; NIKITCZUK, J.; MOREL, G.; MAVROIDIS, C.; WEINBERG, B. Torque Control of Electrorheological Fluidic Resistive Actuators for Haptic Vehicular Instrument Controls. *Journal of Dynamic Systems, Measurement, and Control 128*, 2 (2006), 216–226.
274. VOYLES, R.; MORROW, J. D.; KOHSLA, P. K. The shape from motion approach to rapid and precise force/torque sensor calibration. *Journal of Dynamic Systems, Measurement and Control 119*, 2 (1997), 229–235.
275. VSRHELYI, G.; DM, M.; VZSONYI, .; BRSONY, I. A. Effects of the elastic cover on tactile sensor arrays. *Sensors and actuators. A, Physical 132*, 1 (2006), 245–251.
276. WAGNER, C.; LEDERMAN, S.; HOWE, R. A tactile shape display using RC servomotors. In *Haptic Interfaces for Virtual Environment and Teleoperator Systems, 2002. HAPTICS 2002. Proceedings. 10th Symposium on* (Orlando, FL ISBN: 0-7695-1489-8, 24252002), pp. 354 – 355.
277. WAHED, A. K. The characteristics of a homogeneous electrorheological fluid in dynamic squeeze. In *Actuator 2004: 9th International Conference on New Actuators* (Bremen, 2004), HVG Hanseatische Veranst.-GmbH, Div. Messe Bremen, pp. 605 – 608.
278. WALLASCHEK, J. A. Contact mechanics of piezoelectric ultrasonic motors. *Smart Materials and Structures*, 3 (1998), 369 – 381.
279. Piezoaktoren in ihrer Anwendung.
280. WEAVER, W.; TIMOSENKO, S. P.; YOUNG, D. H. *Vibration problems in engineering*, 5. ed ed. Wiley, New York [u.a.], 1990.
281. WEINBERG, B.; NIKITCZUK, J.; FISCH, A.; MAVROIDIS, C. Development of electrorheological fluidic resistive actuators for haptic vehicular instrument controls. *Smart Materials and Structures 14*, 6 (2005), 1107 – 1119.
282. WEISSMANTEL, H.; HOPPACH, E. Elektrische Kleinantriebe - Skriptum zur Vorlesung. 2002.
283. WEN, W.; HUANG, X.; SHENG, P. Particle size scaling of the giant electrorheological effect. *Applied Physics Letters 85*, 2 (2004), 299–301.
284. WERTHSCHUETZKY, R. *Mess- und Sensortechnik - Band II: Wirkprinzipien von Primärsensoren*. Skriptum, Technische Universität Darmstadt, 2007.
285. WERTHSCHUETZKY, R. *Mess- und Sensortechnik - Band III: Analoge Sensorelektronik*. Skriptum, Technische Universität Darmstadt, Darmstadt, 2007.
286. WIKER, S. F.; HERSHKOWITZ, E.; ZIK, J. Teleoperator Comfort and Psychometric Stability: Criteria for Limiting Master-Controller Forces of Operation and Feedback During Telemanipulation. *Proceedings of NASA Conference on Space Telerobotics, Vol. 1* (1989), 99–107.
287. WILKINSON, R. S.; FUKAMI, Y. Responses of isolated Golgi tendon organs of cat to sinusoidal stretch. *J Neurophysiol 49*, 4 (1983), 976–88.
288. WINFIELD, L.; GLASSMIRE, J.; COLGATE, J. E.; PESHKIN, M. T-PaD: Tactile Pattern Display through Variable Friction Reduction. In *EuroHaptics Conference, 2007 and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems. World Haptics 2007. Second Joint* (Tsukaba, 2007), pp. 421 – 426.
289. WOLF, U. *Aristoteles' "Nikomachische Ethik"*, 2 ed. Wiss. Buchges., [Abt. Verl.], Darmstadt, 2007.
290. YAMAMOTO, A.; ISHII, T.; HIGUCHI, T. Electrostatic tactile display for presenting surface roughness sensation. *Industrial Technology, 2003 IEEE International Conference on 2* (2003), 680 – 684 Vol.2.
291. YAMAOKA, M.; YAMAMOTO, AKIO, H. T. Basic Analysis of Stickiness Sensation for Tactile Displays. In *Eurohaptics Conference* (Madrid, 2008), Springer.
292. YEONGMI, K.; OAKLEY, I.; JEHA, R. Design and Psychophysical Evaluation of Pneumatic Tactile Display. In *SICE-ICASE, 2006. International Joint Conference* (Busan ISBN: 89-950038-5-5, 2006), pp. 1933 – 1938.
293. YOSHIKAWA, T.; ICHINOO, Y. Impedance identification of human fingers using virtual task environment. *Intelligent Robots and Systems, 2003. (IROS 2003). Proceedings. 2003 IEEE/RSJ International Conference on 4* (2003), 3094 – 3099 vol.3.

294. YU, O.; ZHARRI. *An exact mathematical model of a travelling wave ultrasonic motor*, vol. 1 of *Ultrasonics Symposium, 1994. Proceedings., 1994 IEEE*. Cannes, 1-4 Nov. 1994.
295. ZACHMANN, G. *Virtual Reality in Assembly Simulation — Collision Detection*. PhD thesis, Darmstadt University of Technology, Germany, Department of Computer Science, may 2000.
296. ZADEH, M. H.; WANG, D.; KUBICA, E. Perception-based lossy haptic compression considerations for velocity-based interactions. *Multimedia Systems 13*, 4 (2008), 275–282.
297. ZHU, M. *Contact analysis and mathematical modeling of traveling wave ultrasonic motors*, vol. 51 of *Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on*. June 2004.
298. ZIEMANN, O. *POF-Handbuch: optische Kurzstrecken-Übertragungssysteme*, 2 ed. Springer, Berlin [u.a.], 2007.
299. ZILLES, C. B.; SALISBURY, J. K. A constraint-based god-object method for haptic display. In *IROS '95: Proceedings of the International Conference on Intelligent Robots and Systems* (Washington, DC, USA, 1995), vol. 03, IEEE Computer Society, p. 3146.

"*This page left intentionally blank.*"

Index

A

absolute threshold 40
active 10, 27
active haptic interaction 60, 63
actuation principle 192
actuator design 191
actuator designs, piezoelectric 248
actuator, electrostatic 277
actuator, piezoelectric 243
addressability 23
adhesion 28
admittance controlled, closed-loop 100
admittance controlled, open-loop 98
admittance, mechanical 27, 443
ampere turns 203
anthropometric 11
artificial muscle 292
assistive system, haptic 25

B

bar-shaped resonator 250
basic equations, piezoelectric 244
basic piezoelectric actuator designs 248
being, physical 6
bending moment 324
bending-actuator 194, 266
Bernoulli 325
bimorph, piezoelectric 263
block-commutation 218
Bragg 346
Braille 266
brake, electromagnetic 237

C

C 82, material property 247
capacitive actuator 194
capacitive actuators 279
capacitive principle 192
capacitive sensors 366
charge constant 246
closed-loop admittance controlled 95
closed-loop impedance controlled 95
cluster “always” 122
cluster “kinaesthetic” 118
cluster “omni-dimensional” 121
cluster “surface-tactile” 119
cluster “vibro-directional” 121
cluster “vibro-tactile” 120
code disc 360
π-coefficients 330
coefficient of strain, piezoelectric 245
coefficient of tension, piezoelectric 245
coefficients, piezoelectric 245
collision detection
  dynamic 428
  static 428
commutated, electronic 216
commutation 216
compression 380
conductor 201
constraint 183
contact grasp 65
control design
    for Haptic Systems 152
control law design 156
control quality 132
controller, haptic 22
Coulomb’s law 277
coupling factor, piezoelectric 247
Curie temperature 247
current source, analog 223

customer, experiments with 122

D

D/A converter 221
DC-drive 215
DC-motor 193
dead-man-switch 64
Delta.3 189
Delta-Robot 170
Denavit-Hartenberg-parameter 169, 177, 178, 181
design piezoelectric actuators 258
design process 11
designs of DEA 289
determinant 172
dielectric elastomer actuator (DEA) 285, 287
difference threshold 40
Difference-Limen 40
differentiation of signals 369
digital to analog converter 221
direct current magnetic field 207
direct kinematics 169, 190
direct method by LYAPUNOV 148
discretization 87
disturbance compensation 158
DL 40
DoF 391
dots-per-inch 359
DPI 359
driver electronics, electrodynamic 219
dynamic 37
dynamic behavior 132
dynamic collision detection 428
dynamics 46

E

EC-drive 216
EC-motor 193
ECP 382
effect, piezoelectric 243
elastic constant 245
elasto-mechanics 322
electric field 277
electric motor 193
electrical time constant 212
electro-rheological fluid 292
electrochemical principle 192
electrodynamic principle 192
electromagnetic actuators 229
electromagnetic brake 237
electromagnetic principle 192
electromechanical network 261
electronic-commutated 216
electrostatic actuators 277
energy density, magnetic 208
energy, magnetic 229
enhanced parallel port 382
EPP 382
equations traveling wave motor 255
Ethernet 384
event 118
Event-based haptics 378
evolutionary algorithm 174
experiments 122
external supply rate 150
extreme dead times 379

F

FA-I 36
FA-II 37
Fabry-Pèrot 344
Falcon 167
fast-adapting 36
Fechner's law 44
feedforward 159
feedforward filter function 159
FFB 24
fiber bragg grating 346
field driven actuator 277
field plates 364
field response 212
field strength, magnetic 203
filling-factor 203
FIP 82
FireWire 383
flow mode 295
flux 203
flux density, magnetic 203
flux, magnetic 203
foil sensor 334, 339
force sensing resistor 334
force, magnetomotive 203
force-feedback 24
force-impression 82
forward kinematics 169
forward kinematics for serial mechanisms 177
frequency domain analysis 89
frequency-dependency 46, 48
friction 28
fully parallel 183
functional requirements 124

G

Gaussian distribution 45
gear-DOF 183
gears 197
gearwheel 198
Grübler formula 182, 186
gradient approach 174
Grasps 64
gray scale sensor 361

H

H-bridge 221
hall sensors 364
Hammerstein-model 140
HapKeys 226
haptic assistive system 25
haptic controller 22
haptic device 22
haptic device, active 27
haptic device, passive 27
haptic display 24
haptic interaction 23, 117
haptic interface 24
haptic loop 127
haptic manipulator 24
haptic marker 24
haptic rendering 389
haptic shape 28
haptic simulator 25
haptic surface properties 28
haptic texture 28
haptics 21
hardware in the loop 384
harmonic balance 145
hexapod 166
HID 382, 393
high-grade parallel 183
HIL 384
human-interface-device 382
hybrid kinematic 166, 188
hydraulic 194

I

identical condition 183
IEEE1394 383
impedance 63
impedance controlled, closed loop 97
impedance controlled, open-loop 96
impedance coupling 62, 64
impedance, mechanical 27, 443
Induction 211
input strictly passive 150
integral criterion 157
integration of signals 369
intensity 341
interaction metaphor 388, 403
interaction, haptic 23
interface 17
interface, haptic 24
inverse kinematics 169, 187, 190
iron-less rotor 213

J

Jacobian-matrix 169
Jacobian matrix, direct 172
Jacobian matrix, inverse 172
Jacobian matrix, rank 173
Jacobian matrix, singular values 174
JND 40, 319
joint, driven 166
joint, linear 175
joint, passive 166
joint, rotatory 176
joint, un-driven 166
Just-Noteable-Difference 40
Just-Noticeable-Difference 40

K

k-factor 326, 353
kinaesthetic 321
kinaesthetic perception 10
kinaesthetic sensors 37
kinaesthetics 21
kinematic dependencies 170
kinematic design, complete process 188
kinematic scheme 184
kinematics 166

L

level of parallelism 183
linear SISO-systems 132
Lissajusfigure 251
LM 43
local haptic model 377
longitudinal actuator 246
longitudinal effect, magnetic 231
longitudinal effect, piezoelectric 244
Lorentz-force 199
Lossless 150
Lyapunov 145
Lyapunov functions 148

M

magn. resistance 203
magnetic circuits 203, 232
magnetic cross section 232
magnetic dependent resistors 364
magnetic energy 229, 234
magnetic field strength 203
magnetic field, direc current 207
magnetic flux 203
magnetic flux density 203, 205
magnetic resistance 203
magneto-rheological-fluid 297
magnetomotive force 203
magnetorheological principle 192
manipulator, haptic 24
marker, haptic 24
masking 41
material properties, piezoelectric 247
materials, piezoelectric 247
mechanical admittance 27
mechanical commutation 215
mechanical impedance 27
mechanically commutated 217
mechanism 166, 170
mechanism, parallel 182, 186
mechanism, serial 175
Meissner 36
Merkel 37
mice-sensor 362
micro frame 382
micro-bending sensor 343
model time 397
motor cortex 36
moving coils, electrodynamic 213
moving magnet 216
multi-input-multi-output-systems 136
multimodality 40

N

natural behavior 389
natural interaction 388
neurobiology 19
nociceptive 21
nominal load 314, 354
nonlinear stability analysis 145
nonlinear system description 139
Nyquist´s stability criterion 142

O

observer based
state space control 160
open-loop admittance controlled 95
open-loop impedance controlled 95
optical position sensors 360
output equation 137
output strictly passive 150
overshoot 154

P

Pacinian 37
parallel kinematic 165, 166, 188
parallel port 382
parallel-plate capacitor 277
partly parallel 183
passive 10, 27
passive haptic device 90
passive haptic interaction 60
passivity 145
passivity, control engineering 63
percentile 45
percentiles 12
Perception 80
perception, kinaesthetic 84
perception, tactile 85
permanent magnet 206
permanent magnets 234
permeability 203
permeability number 234
permittivity 203
PHANTOM Omni haptic device 176
phase plane analysis 145
photo-elastic effect 340
photodiode 361
phototransistor 361
physical-interface-device 382
PID 382
PID-Control 156
piezoelectric actuators 243
piezoelectric actuators, design 258
piezoelectric basic equations 244
piezoelectric bimorph 263
piezoelectric coefficient of strain 245
piezoelectric coefficient of tension 245
piezoelectric coefficients 245
piezoelectric coupling factor 247
piezoelectric effect 243
piezoelectric equation 246
piezoelectric longitudinal effect 244
piezoelectric material properties 247
piezoelectric materials 247
piezoelectric motor 194
piezoelectric principle 192
piezoelectric sensors 347

piezoelectric shear effect 244
piezoelectric special designs 250
piezoelectric stack 194
piezoelectric stepper motors 257
piezoelectric transversal effect 244
piezoelectrical Bimorph 266
piezoelelectric shear effect 257
Pixel 391
plunger type magnet 238
plunger-type magnet 194
pneumatic 194
Popov inequality 146
Popov plot 146
Popov-criterion 145
positioning system 166
power grasp 65
power grasps 70
power law 44
power loss, electrodynamic 201
precision grasp 65
precision grasps 71
propagation 395
psychophysics 19, 40
PT_1-System 135
PT_2-System 137
puls-width-modulation 221
PVDF, material property 247
PWM 221
PZT-4, material property 247
PZT-5a, material property 247

Q

quadrant controllers 219
Quarz, Kristallstruktur 243
Quarz, material property 247

R

RA 36
rapid-adapting 36
rare earth 206
real-time 397
reflection light switches 362
relative resistivity change 326
reluctance 203
reluctance drives 237
reluctance effect 229
reluctance effect, magnetic 231
remanence flux density 208
rendering 389
requirements 124, 319, 352
resistance, magnetic 203
resistivity change 330
resolution 23
resonance principle 350
resonance-actuator 193
resonator, bar-shaped 250
resonator, circular 254
reverse kinematics 169
robot 166
robot arm 166
root locus method 142, 157
roughness 29
Routh Hurwitz criterion 142
Ruffini 37

S

SA-I 37
SA-II 37
Sashida 254
SAW sensors 351
SCARA-robot 167
self-supportive 215
semiconductor 328
senses 5
sensor-less commutation 218
sensors, kinaesthetic 37
sensors, tactile 36
serial kinematic 165, 166, 188
serial port 381
servo-drive 216
shaker 193
shape 117
shape, haptic 28
shape-memory alloy 192
shape-memory wire 194
shear effect, piezoelectric 244, 257
shear mode 294
silicon sensors 331
simulator system 374
simulator, haptic 25
singular position 172
singularity 172
sinus-commutation 218
slow-adapting 37
somatotrope 36
spatial masking 41
special designs, piezoelectric 250
specifications 124
squeeze mode 295
stability analysis 141
standing wave 250
state feedback control 159
state space control 159
state space representation 136

state space vector 138
state strictly passive 150
static collision detection 428
stepper motor 236
stepper motors, piezoelectric 257
stepper-motor 194
Stevens 44
Steward-Gough-platform 166
strain gauge 327
stress 324
strictly passive 150
Successivenes Limen 43
surface adhesion 28
surface friction 28
surface micro-machining 338
surface properties, haptic 28
surface-wave actuators 195
surgery 14
system description 132
System Engineering 95
system equation 136
system specification 124
system stability 131
system strukturing 152
system transfer function 141

T

tactile 21, 317, 319, 332, 337, 350
tactile perception 10
tactile sensor 36
tactile systems 290, 298
tacton 16, 29
tactor 24
telemanipulation 14
telemanipulation system 25, 374
telemanipulator 25
temperature, monitor 224
temporal masking 41
texture 117, 317
Texture Explorer 267
texture, haptic 28
thermal principle 192
thermoceptive 21
threshold 40
threshold, difference 40
time constant, electrical 212
Tool Center Point, TCP 168
topology 166, 189
total reflection 341
transfer characteristics 170, 173
transfer characteristics, optimization 174
transmission chain 374
transmission-ratio 197
transparency 27
transversal actuator 246
transversal effect 229
transversal effect, electromagnetic 230
transversal effect, piezoelectric 244
traveling wave motor, equations 255
traveling wave motor, linear 254
travelling wave 254
travelling wave motor 273
triangulation 362
two-point threshold 52
two-point-threshold 42

U

Ubi-Pen 268
Uchinomotor, Uchino 255
Ultraschallaktoren 250
ultrasonic sensors 365
ultrasonic-actuator 193
universal serial bus 382
upper cut-off frequency 354
USB 382
user 22, 29
user-impedance 27

V

vibrotactile display 267
virtual work 173
viscoelastic material performance 316
voice-coil-actuator 193
voigtsche notation 245

W

wave, standing 250
wave, traveling 254
Weber's law 43
Wheatstone 326
Wiener-Model 139
wire 201

Z

Z 27
Z-width 27